PCT

INTERNATIONAL APPLICATION PUBLISHE

(51) International Patent Classification 6 : G01N 33/53, C12Q 1/70, C12N 5/10, 7/01, C07K 7/00, 14/00, A61K 38/03	**A1**

WO 9603649A1

(43) International Publicati n Date: 8 February 1996 (08.02.96)

(21) International Application Number: PCT/US95/09382

(22) International Filing Date: 24 July 1995 (24.07.95)

(30) Priority Data:
278,865   22 July 1994 (22.07.94)   US
483,555   7 June 1995 (07.06.95)   US

(71) Applicant: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; Office of Technology Develop-ment, CB #4100, 302 Bynum Hall, Chapel Hill, NC 27599-4100 (US).

(72) Inventors: SPARKS, Andrew, B.; 201 Blue Ridge, Carrboro, NC 27510 (US). KAY, Brian, K.; 18 Wysteria Way, Chapel Hill, NC 27514 (US). THORN, Judith, M.; Apartment 18-G, 200 Barnes, Carrboro, NC 27510 (US). QUILLIAM, Lawrence, A.; Apartment 5, 206 Pinegate Circle, Chapel Hill, NC 27514 (US). DER, Channing, J.; 101 Fieldstone Court, Chapel Hill, NC 27514 (US).

(74) Agents: MISROCK, S., Leslie et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.

(54) Title: Src SH3 BINDING PEPTIDES AND METHODS OF ISOLATING AND USING SAME

(57) Abstract

Peptides having general and specific binding affinities for the Src homology region 3 (SH3) domains of proteins are disclosed in the present invention. In particular, SH3 binding peptides have been isolated from three phage-displayed random peptide libraries which had been screened for isolates that bind to bacterial fusion proteins comprising SH3 and glutathione S-transferase (GST). Preferred peptides are disclosed which comprise a core 7-mer sequence (preferably, a consensus motif) and two or more, preferably at least six, additional amino acid residues flanking the core sequence, for a total length of 9, preferably at least 13, amino acid residues and no more than about 45 amino acid residues. Such peptides manifest preferential binding affinities for certain SH3 domains. The preferred peptides exhibit specific binding affinities for the Src-family of proteins. *In vitro* and *in vivo* results are presented which demonstrate the biochemical activity of such peptides.

Src SH3 BINDING PEPTIDES AND
METHODS OF ISOLATING AND USING SAME

1.   Field of the Invention

5   The present invention relates to SH3 binding peptides
having a broad range of binding specificities. That is, certain
members of the SH3 binding peptides disclosed bind with
approximately the same facility with SH3 domains derived from
different SH3 domain-containing proteins. Other members, in
10   contrast, bind with a much greater degree of affinity for
specific SH3 domains. The SH3 binding peptides are obtained from
random peptide libraries that are also phage-displayed. Methods
are described of obtaining the phage clones that bind to the SH3
domain targets and of determining their relevant nucleotide
15   sequences and consequent primary amino acid sequence of the
binding peptides. The resulting SH3 binding proteins are useful
in a number of ways, including, but not limited to, providing a
method of modulating signal transduction pathways at the cellular
level, of modulating oncogenic protein activity or of providing
20   lead compounds for development of drugs with the ability to
modulate broad classes, as well as specific classes, of proteins
involved in signal transduction.

2.   Background of the Invention
2.1. Src and the SH3 Domain

25   Among a number of proteins involved in eukaryotic
cell signaling, there is a common sequence motif called the SH3
domain. It is 50-70 amino acids in length, moderately conserved
in primary structure, and can be present from one to several
times in a large number of proteins involved in signal
30   transduction and in cytoskeletal proteins.
        The protein pp60c-src represents a family of at least nine
non-receptor protein tyrosine kinases (NR-PTKs). Members of this
family share an overall structural organization comprising a

series of catalytic and non-catalytic domains. In Src, a 14-
amino-acid myristylation ·signal resides at the extreme amino-
terminus, and is followed by a unique region that is not highly
conserved among family members. Following this region are two
5   highly conserved 60- and 100-amino-acid regions, the Src homology
(SH) domains 3 and 2, respectively. SH2 and SH3 domains have
been shown to play an important role in mediating protein-protein
interactions in a variety of signaling pathways. Koch, C.A., et
al., in Science (1991) 252:668-74. The carboxy-terminal half of
10   Src contains the PTK catalytic domain, as well as a negative
regulatory tyrosine (Y527) near the carboxy terminus.
Phosphorylation of this residue (e.g., by Csk) results in the
inhibition of PTK activity. Cooper, J.A., et al., in Science
(1986) 231:1431-1434. Mutation of Y527->F generates forms of Src
15   with increased PTK and oncogenic activity. Cartwright, C.A., et
al., in Cell (1987) 49:83-91; Kmiecik, T.E., et al., in Cell
(1987) 49:65-73; and Piwicna-Worms, H., et al., in Cell (1987)
75-82.
       The fact that some mutations which result in increased Src
20   PTK and transforming activity map to the Src SH2 (Seidel-Dugan,
C., et al., in Mol. Cell. Biol. (1992) 12:1835-45; and Hirai, H.
and Varmus, H.E. in Mol. Cell. Biol. (1990) 10:1307-1318) and SH3
domains (Seidel-Dugan, C., et al., supra; Hirai, H. and Varmus,
H.E., supra; Superti-Furga, G., et al., in Embo. J. (1993)
25   12:2625-34; and Potts, W.M., et al., in Oncogene Res. (1988)
3:343-355) suggests a negative regulatory role for these domains.
That phosphotyrosine residues within specific sequence contexts
represent high affinity ligands for SH2 domains suggests a model
in which the SH2 domain participates in Y527-mediated inhibition
30   of PTK activity by binding phosphorylated Y527, thereby locking
the kinase domain in an inactive configuration. Matsuda, M.,
Mayer, B.J., et al., in Science (1990) 248:1537-1539. This model
is supported by the observation that phosphopeptides
corresponding to the carboxy-terminal tail of Src bind active,
35   but not inactive, variants of Src. Roussel, R.R., et al., in

<u>Proc. Natl. Acad. Sci. U S A</u> (1991) 88:10696-700; and Liu, X., et al., in <u>Oncogene</u> (1993) 8:1119-1126.

The mechanism of SH3-mediated inhibition of Src PTK activity remains unclear. There is evidence that pY527-mediated

5   inhibition of Src PTK activity involves the SH3 domain as well as the SH2 domain. Okada, M., Howell, et al., in <u>J. Biol. Chem.</u> (1993) 268:18070-5; Murphy, S.M., et al., in <u>Mol. Cell. Biol.</u> (1993) 13:5290-300; and Superti-Furga, G., et al., <i>supra</i>. Although these effects are thought to be a consequence of SH3-

10   mediated protein-protein interactions, precisely how the Src SH3 domain exerts its negative regulatory effect is unclear. Identification of high affinity ligands for the Src SH3 domain could help resolve these issues.

2.2. Protein Tyrosine Kinases and The Immune Response

15   Src-related tyrosine kinases are expressed in a variety of cell types including those of the immune system (lymphocytes, T cells, B cells, and natural killer cells) and the central nervous system (neural cells, neurons, oligodendrocytes, parts of the cerebellum, and the like). Umemori, H. et al., in <u>Brain Res.</u>

20   <u>Mol. Brain Res.</u> (1992) Dec. 16(3-4):303-310. Their presence in these cells and tissues and their interaction with specific cell surface receptors and immunomodulatory proteins (such as T cell antigen receptor, CD14, CD2, CD4, CD40 or CD45) suggest that these kinases serve an important role in the signalling pathways

25   of not only the central nervous system but of the immune system, as well. See, e.g., Ren, C.L. et al., in <u>J. Exp. Med.</u> (1994) 179(2):673-680 (signal transduction via CD40 involves activation of Lyn kinase); Donovan, J.A. and Koretzky, G.A., in <u>J. Am. Soc.</u> <u>Nephrol.</u> (1993) 4(4):976-985 (CD45, the immune response, and

30   regulation of Lck and Fyn kinases); and Carmo, A.M. et al., in <u>Eur. J. Immunol.</u> (1993) 23(9):2196-2201 (physical association of the cytoplasmic domain of CD2 with p56lck and p59fyn).

For instance, mice with disruptions in their Src-like genes, Hck and Fgr, possess macrophages with impaired phagocytic

- 3 -

activity or exhibit a novel immunodeficiency characterized by an
increased susceptibility to infection with Listeria
monocytogenes. Lowell, C.A. et al., in <u>Genes Dev.</u> (1994)
8(4):387-398. Also, it has been shown that bacterial
5   lipopolysaccharide (LPS) activates CD14-associated p56lyn,
p68hck, and p59c-fgr, while inducing the production of
lymphokines, such as TNF-alpha, IL-1, IL-6, and IL-8. Inhibition
of the protein tyrosine kinases blocks production of TNF-alpha
and IL-1.

10   **2.3. SH3 Binding Peptides**

As mentioned above, it has long been suspected that SH3
domains are sites of protein-protein interaction, but it has been
unclear what SH3 domains actually bind. Efforts to identify
ligands for SH3 domains have led to the characterization of a
15   number of SH3-binding proteins, including 3BP1 and 2 (Ren, R.,
Mayer, et al., in <u>Science</u> (1993) 259:1157-61), SOS (Olivier,
J.P., et al., in <u>Cell</u> (1993) 73:179-91; and Rozakis-Adcock, M.,
et al., in <u>Nature</u> (1993) 363:83-5), p85 PI-3' Kinase (Xingquan,
L., et al., in <u>Mol. Cell. Biol.</u> (1993) 13:5225-5232), dynamin
20   (Gout, I., et al., in <u>Cell</u> (1993) 75:25-36), AFAP-110 (Flynn,
D.C., et al., in <u>Mol. Cell. Biol.</u> (1993) 13:7892-7900), and CD42
(Barfod, E.T., et al., in <u>J. Biol. Chem.</u> (1993) 268:26059-26062).
These proteins tend to possess short, proline-rich stretches of
amino acids, some of which have been directly implicated in SH3
25   binding. A variety of consensus sequences have been proposed,
although the similarity among proline-rich regions of different
SH3-binding proteins tends to be fairly low. Also, attempts to
build consensus sequences are likely complicated by the
incorporation of data from proteins that bind different SH3
30   domains.

Thus, Cicchetti, P., et al., in <u>Science</u> (1992) 257:803-806,
published their work relating to the isolation and sequencing of
two naturally-occurring proteins that could be bound *in vitro* by
the SH3 domain of the *abl* oncogene product. These workers found

- 4 -

that SH3 domains bind short, proline-rich regions of such
proteins. Subsequently, this same group disclosed further
results (Ren, R. et al., *supra*) in which the SH3 binding sites of
the SH3 binding proteins were localized to "a nine- or ten-amino
5   acid stretch rich in proline residues." A consensus sequence
incorporating the features of the SH3 binding sites of four SH3
binding proteins was proposed: XPXXPPPΨXP (SEQ ID NO:1), wherein
X indicates a position in the amino acid sequence which is not
conserved among the four SH3 binding proteins, P represents
10   proline, and Ψ indicates a hydrophobic amino acid residue, such
as P or L.

The screening of complex random peptide libraries has been
used to identify peptide epitopes for monoclonal (Scott, J.K. and
Smith, G.P. in Science (1990) 249:386-390) and polyclonal (Kay,
15   B.K., et al., in Gene (1993) 128:59-65) antibodies, as well as
peptide ligands for a variety of proteins, including streptavidin
(Devlin, J.J., et al., in Science (1990) 249:404-406; and Lam,
K., et al., in Nature (1991) 354:82-84), the endoplasmic
reticulum chaperone BiP (Blond-Elguindi, S., et al., in Cell
20   (1993) 75:717-728), and CaM (Dedman, J.R., et al., in J. Biol.
Chem. (1993) 268:23025-23030).

Recently, Chen, J.K. et al., in J. Am. Chem. Soc. (1993)
115:12591-12592, described ligands for the SH3 domain of
phosphatidylinositol 3-kinase (PI-3' Kinase) which were isolated
25   from a biased combinatorial library. A "biased" library is to be
distinguished from a "random" library in that the amino acid
residue at certain positions of the synthetic peptide are fixed,
i.e., not allowed to vary in a random fashion. Indeed, as stated
by these research workers, screening of a "random" combinatorial
30   library failed to yield suitable ligands for a PI-3' Kinase SH3
domain probe. The binding affinities of these unsuitable ligands
was described as weak, >100 μM, based on dissociation constants
measured by the Biosensor System (BIAcore).

- 5 -

More recently, Yu, et al. (Yu, H., et al., in Cell (1994) 76:933-945) used a "biased" synthetic peptide library of the form XXXPPXPXX (SEQ ID NO:2), wherein X represents any amino acid other than cysteine, to identify a series of peptides which bind

5   the Src and PI-3' Kinase SH3 domains. The bias was accomplished by fixing the proline residues at the specific amino acid positions indicated for the "random" peptide. As stated previously, without this bias, the technique disclosed fails to identify any SH3 domain-binding peptides.

10.   A consensus sequence, based on 13 binding peptides was suggested: RXLPPRPXX (SEQ ID NO:3), where X tends to be a basic residue (like R, K or H). The binding affinities of several SH3 binding peptides were disclosed as ranging from 8.7 to 30 μM. A "composite" peptide, RKLPPRPRR (SEQ ID NO:4), was reported to

15   have a binding affinity of 7.6 μM. This value compares favorably to the binding affinity of the peptide, VPPPVPPRRR (SEQ ID NO:5), to the N-terminal SH3 domain of Grb2. See, Kraulis, P.J. J. Appl. Crystallogr. (1991) 24:946. Recognizing the limitations of their technique, Chen and co-workers, supra, stated that their

20   results "illustrate the utility of biased combinatorial libraries for ligand discovery in systems where there is some general knowledge of the ligand-binding characteristics of the receptor" (emphasis added).

       Yu and co-workers, supra, further described an SH3 binding

25   site consensus sequence, Xp∅PpXP (SEQ ID NO:6), wherein X represents non-conserved residues, ∅ represents hydrophobic residues, P is proline, and p represents residues that tend to be proline. A consensus motif of RXLPPRPXX (SEQ ID NO:7), where X represents any amino acid other than cysteine, was proposed for

30   ligands of PI-3' Kinase SH3 domain. A consensus motif of RXLPPLPRϕ (SEQ ID NO:8), where ϕ represents hydrophobic residues, was proposed for ligands of Src SH3 domain. Still, the dissociation constants reported for the 9-mer peptides ranged only from about 8-70 μM and selectivity between one type of SH3

- 6 -

domain and another was relatively poor, the K_Ds differing by only
about a factor of four.   .

5   Hence, there remains a need to develop techniques for the
identification of Src SH3 binding peptides which do not rely on
such "biased" combinatorial peptide libraries that are limited to
a partially predetermined set of amino acid sequences. Indeed,
the isolation of SH3 binding peptides from a "random" peptide
library has not been achieved successfully before now.
Furthermore, particular peptides having much greater binding
10   affinities, whether general or more selective binding for
specific SH3 domains, remain to be identified. Binding peptides
specific for particular SH3 domains are useful, for example, in
modulating the activity of a particular SH3 domain-containing
protein, while leaving others bearing an SH3 domain unaffected.
15   Still, the more promiscuous general binding peptides are useful
for the modulation of a broad spectrum of SH3 domain-containing
proteins.

         The present invention relates to such SH3 binding peptides,
methods for their identification, and compositions comprising
20   same. In particular, peptides comprising particular sequences of
amino acid residues are disclosed which were isolated from random
peptide libraries. In the present invention, clones were
isolated from a phage-displayed random peptide library which
exhibited strong binding affinities for SH3 domain-containing
25   protein targets. Some of these protein targets, include Abl,
Src, Grb2, PLC-δ, PLC-γ, Ras GAP, Nck, and p85 PI-3' Kinase.
From the nucleotide sequence of the binding phage, the amino acid
sequence of the peptide inserts has been deduced. Synthetic
peptides having the desired amino acid sequences are shown to
30   bind the SH3 domain of the target proteins. In particular,
synthetic peptides combining a core consensus sequence and
additional amino acid residues flanking the core sequence are
especially effective at binding to particular target protein SH3
domains. The SH3 binding peptides disclosed herein can be
35   utilized in a number of ways, including the potential modulation

of oncogenic protein activity *in vivo*. These peptides also serve
as useful leads in the production of peptidomimetic drugs that
modulate a large class of proteins involved in signal
transduction pathways and oncogenesis.

5   3.   <u>Summary of the Invention</u>

              Accordingly, three phage-displayed random peptide
libraries were screened for isolates that bind to bacterial
fusion proteins consisting of the Src homology region 3 (SH3) and
glutathione S-transferase (GST). DNA sequencing of the isolates
10   showed that they contained sequences that resemble the consensus
motif, RPLPPLP (SEQ ID NO:9), within their 8, 22, or 36 amino
acid long random regions. When peptides were synthesized
corresponding to the pIII inserts of the SH3-binding phage, they
bound to the GST fusions of the SH3 domains of Src and the Src-
15   related proteins, such as Yes, but not of Grb2, Crk, Abl, or
PLCγ1. The synthesized peptides bind quite well to the Src SH3
domain and act as potent competitors of natural Src SH3
interactions in cell lysates. For instance, these peptides can
compete with radiolabelled proteins from cell lysates in binding
20   to immobilized Src-GST, with an apparent IC_{50} of 1-10 μM. When a
peptide, bearing the consensus sequence RPLPPLP (SEQ ID NO:9) was
injected into *Xenopus laevis* oocytes, it accelerated the rate of
progesterone-induced maturation. These results demonstrate the
utility of phage-displayed random peptide libraries in
25   identifying SH3-binding peptide sequences and that such
identified peptides exhibit both *in vivo* and *in vitro* biological
activity.

              Thus, it is an object of the present invention to provide
peptides having at least nine and up to forty-five amino acid
30   residues, including an amino acid sequence of the formula,
R-2-L-P-5-6-P-8-9 (SEQ ID NO:10), positioned anywhere along the
peptide, in which each number represents an amino acid residue,
such that 2 represents any amino acid residue except cysteine, 5
and 6 each represents a hydrophobic amino acid residue, 8

- 8 -

represents any amino acid residue except cysteine, and 9
represents a hydrophilic amino acid residue except cysteine, each
letter being the standard one-letter symbol for the corresponding
amino acid, said peptide exhibiting a binding affinity for the
5   SH3 domain of Src, provided that said peptide is not R-P-L-P-P-L-
P-T-S (SEQ ID NO:11). In a particular embodiment of the present
invention, the peptides also exhibit a binding affinity for the
SH3 domain of Src-related proteins, including Yes, Fyn, Lyn, Lck,
Hck and Fgr.

10   The present invention also contemplates SH3 domain-binding
peptides that further comprise a C-terminal-flanking amino acid
sequence of the formula 10, 10-11, 10-11-12, 10-11-12-13, (SEQ ID
NO:12) or 10-11-12-13-14 (SEQ ID NO:13), in which each number
represents any amino acid residue except cysteine, such that 10
15   is bound to 9 by a peptide bond. Furthermore, peptides are also
provided which further comprise an N-terminal-flanking amino acid
sequence of the formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ
ID NO:14) in which each number represents any amino acid residue
except cysteine, such that 1' is bound to R by a peptide bond.

20   Thus, in a particular embodiment, a peptide is disclosed
having at least thirteen and up to forty-five amino acid
residues, including an amino acid sequence of the formula, 3'-2'-
1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned anywhere along
the peptide, in which each number represents an amino acid
25   residue, such that 3', 2', 1', 2, 8, and 10 each represents any
amino acid residue except cysteine, 5 and 6 each represents a
hydrophobic amino acid residue, and 9 represents a hydrophilic
amino acid residue except cysteine, each letter being the
standard one-letter symbol for the corresponding amino acid, said
30   peptide exhibiting a binding affinity for the SH3 domain of Src.

          The present invention also seeks to provide new consensus
sequences or motifs that reflect variations in SH3 domain binding
selectivities or specificities. The present invention also
contemplates conjugates of the SH3 binding peptides and a second
35   molecule or chemical moiety. This second molecule may be any

- 9 -

desired substance whose delivery to the region of the SH3 domain
of a particular protein (or cell containing the protein) is
sought. Possible target cells include, but are not limited to,
neural cells, immune cells (e.g., T cells, B cells, natural

5   killer cells, and the like), osteoclasts, platelets, epidermal
cells, and the like, which cells express Src, Src-related
proteins, and potentially, other SH3 domain-containing proteins.
In this manner, the modulation of the biological activity of
proteins bearing an SH3 domain can be accomplished.

10   Other methods and compositions consistent with the
objectives of the present invention are likewise disclosed. In
particular, a method is disclosed of modulating the activity of
Src or Src-related proteins comprising administering a
composition comprising an effective amount of a peptide of the

15   present invention and a carrier, preferably a pharmaceutically
acceptable carrier. In a specific embodiment, the contemplated
method results in the inhibition of the activity of Src or Src-
related proteins. Alternatively, the method is effective to
activate Src or Src-related proteins.

20   In yet another embodiment, a method is disclosed of
identifying a peptide having a region that binds to an SH3 domain
comprising: (a) providing an immobilized target protein
comprising an SH3 domain; (b) incubating the immobilized target
protein with an aliquot taken from a random peptide library; (c)

25   washing unbound library peptides from the immobilized target
protein; (d) recovering the peptide bound to the immobilized
target protein; and (e) determining the primary sequence of the
SH3 domain-binding peptide.
            Moreover, a method is disclosed of imaging cells, tissues,

30   and organs in which Src or Src-related proteins are expressed,
which comprises administering an effective amount of a
composition comprising an SH3 domain-binding peptide conjugated
to detectable label or an imaging agent.
            Other objectives of the present invention will become

35   apparent to one of ordinary skill in the art after consideration

- 10 -

of the above disclosure and the following detailed description of
the preferred embodiments.

4.   Brief Description of the Figures

FIG. 1 illustrates a scheme for the generation of a
5   random 36 amino acid peptide library (TSAR-9; e.g., SEQ ID
NO:16). Oligonucleotides were synthesized (SEQ ID NOS:17-18),
converted into double-stranded DNA, cleaved with restriction
enzymes (SEQ ID NOS:19-20), and cloned into the M13 vector, m663.
The random peptide region encoded by the oligonucleotides is
10 shown in the box (SEQ ID NO:16) and is situated at the N-terminus
of mature protein III (SEQ ID NO:21).

FIG. 2 illustrates a scheme for the generation of a random
22 amino acid peptide library (TSAR-12; e.g., SEQ ID NO:23).
Oligonucleotides were synthesized (SEQ ID NOS:24-25), converted
15 into double-stranded DNA, cleaved with restriction enzymes (SEQ
ID NOS:26-27), and cloned into the M13 vector, m663. The random
peptide region encoded by the oligonucleotides is shown in the
box (SEQ ID NO:23) and is situated at the N-terminus of mature
protein III (SEQ ID NO:28).

20 FIG. 3 illustrates a scheme for the generation of a random 8
amino acid peptide library (R8C; SEQ ID NO:30). Oligonucleotides
were synthesized (SEQ ID NOS:31-32), converted into double-
stranded DNA, cleaved with restriction enzymes (SEQ ID NOS:33-
34), and cloned into the M13 vector, m663. The random peptide
25 region (SEQ ID NO:30) is flanked by cysteine residues and is
situated at the N-terminus of mature protein III (SEQ ID NO:35).

FIG. 4 illustrates the possible origin of one class of
double-insert R8C recombinants (e.g., encoding SEQ ID NO:36).
Double-stranded oligonucleotides (e.g., SEQ ID NO:37) may have
30 ligated in a head-to-head fashion at the _Xba_ I site prior to
cloning in the _Xho_ I- _Xba_ I cleaved M13 vector.

FIG. 5 shows a list of random peptide recombinants (SEQ ID
NOS:38-61 and 106) isolated by the method of the present
invention and the displayed peptide sequence. The amino acid

- 11 -

sequences are aligned to highlight the core sequences. The
flanking sequences are shown to the N-terminal and C-terminal
ends of the core sequence.

5   FIG. 6 graphically illustrates the relative binding
affinities of selected phage clones for various SH3 domains. The
results indicate that certain amino acid sequences provide
generic SH3 domain binding, while others can provide greater
selectivity for the SH3 domain of Src. Still other clones
exhibit Src SH3 domain preferential binding.

10   FIG. 7 shows the binding of synthetic peptides (SEQ ID NOS:9
and 62-70) representing Src SH3-selected phage inserts to Src
SH3-GST fusion target (shaded columns) over background GST
binding (unshaded columns) relative to the core peptide RPLPPLP
(SEQ ID NO:9) and proline-rich peptide segments derived from

15   naturally occurring proteins. Bound biotinylated peptide was
detected with streptavidin-alkaline phosphatase ELISA. Each
point was performed in triplicate; average absorbance at 405 nm
is presented. Error bars represent SD.

       FIG. 8 illustrates the relative specificity of selected

20   peptides (SEQ ID NOS:9 and 62-70) for SH3 domains derived from
different proteins. In particular, the binding affinities of the
peptides for the SH3 domains of the following protein fusion
targets were tested: Src SH3-GST, Yes SH3-GST, Grb2-GST, Crk SH3-
GST, Abl SH3-GST, PLCγ1 SH2SH3-GST. Bound biotinylated peptide

25   was detected with streptavidin-alkaline phosphatase. Each point
was performed in triplicate; values are average signal
(absorbance at 405 nm) above GST background, with error bars
representing standard deviation. Hatched bars indicate
saturation of the ELISA signal.

30   FIG. 9 presents the results of competition experiments in
which selected peptides were found to inhibit the binding of
proteins from cell lysates to immobilized Src SH3-GST or Abl SH3-
GST protein fusion targets.

       FIG. 10 presents a graph illustrating the increased rate of

35   progesterone-induced maturation of oocytes injected with an SH3

                              - 12 -

domain-binding peptide, VLKRPLPIPPVTR (SEQ ID NO:64), of the
present invention. Briefly, Stage VI oocyted were prepared and
injected as previously described (see, Kay, B.K., in Methods in
Cell Biol. (1991) 36:663-669). Oocytes were injected with 40 nL

5   of 100 μM test peptide or water. After injection, the oocytes
were placed in 2 μg/mL progesterone (Sigma, St. Louis, MO) and
scored hourly for germinal vesicle breakdown (GVBD).
        FIG. 11 shows the results of fluorescence experiments in
which certain peptides, Panel A = VLKRPLPIPPVTR (SEQ ID NO:64),

10 Panel B = GILAPPVPPRNTR (SEQ ID NO:63), Panel C = RSTPRPLPPLPTTR
(SEQ ID NO:67), of the invention were shown to localize within
cellular compartments thought to contain Src or Src-related
proteins.
        FIG. 12 illustrates a scheme for the generation of a biased

15 peptide library. Oligonucleotides were synthesized, converted
into double-stranded DNA, cleaved with restriction enzymes XhoI
and XbaI, and cloned into the mBAX vector, described further
below in the Examples section. The biased peptide region is
situated at the N-terminus of mature pIII protein.

20   FIG. 13 illustrates the peptide sequence encoded in the mBAX
vector situated at the N-terminus of mature pIII protein.

        5.   Detailed Description of the Preferred Embodiments
              5.1. General Considerations
                    The present invention relates to peptides that

25 exhibit a binding affinity for an SH3 domain, which domain has
been found to be present in a number of physiologically
significant proteins. In particular, peptides are disclosed
which exhibit general binding characteristics to the SH3 domains
found in a group of proteins, including but not limited to Abl,

30 Src, Grb2, PLC-δ, PLC-γ, Ras GAP, Nck, and p85 PI-3' Kinase.
Preferred peptides exhibit selective, if not specific, binding
affinity for the SH3 domain of Src. As described herein, the
peptides of the present invention include a core sequence,
preferably a consensus seqeunce, and additional amino acid

                            - 13 -

residues that flank the core sequence. These peptides, including
the methods for their identification, are described in greater
detail, below.

Thus, in a specific embodiment of the invention, peptides
5   are provided which have at least nine and up to about forty-five
amino acid residues, including an amino acid sequence resembling
the formula,

R-2-L-P-5-6-P-8-9 (SEQ ID NO:10),

positioned anywhere along the peptide. In the above-mentioned
10   formula, each number represents an amino acid residue, such that
2 represents any amino acid residue except cysteine, 5 and 6 each
represents a hydrophobic amino acid residue, 8 represents any
amino acid residue except cysteine, and 9 represents a
hydrophilic amino acid residue except cysteine. Each letter used
15   in the formulas herein represent the standard one-letter symbol
for the corresponding amino acid. When the peptide is a 9-mer,
the peptide R-P-L-P-P-L-P-T-S (SEQ ID NO:11) is excluded. The
peptides of particular interest are those that exhibit a binding
affinity for the SH3 domain of Src and Src-related proteins,
20   including Yes, Fyn, Lyn, Lck, Hck and Fgr. Preferably, the
peptides of the invention exhibit a binding affinity for the SH3
domain of Src, which is at least three-fold, more preferably at
least four-fold, most preferably at least about five-fold greater
than that exhibited by the peptide RPLPPLP (SEQ ID NO:9). In
25   still other embodiments, the peptides exhibit a binding affinity
for the SH3 domain of Src which is at least ten-fold greater than
that exhibited by the peptide RPLPPLP (SEQ ID NO:9).

In specific embodiments, peptides are disclosed in which the
various amino acid residues at the indicated positions may
30   independently have the following preferred identities: 2 is a P,
R, A, L, Q, E or S, more preferably P or R; 5 represents a P, M,
I or L, more preferably P or M; 6 is a P, L, I or V, more
preferably P or L; 8 is a T, R, P, I, N, E, V, S, A, G or L, more
preferably T or R; and 9 is a T, R, S, H or D, more preferably T
35   or R. Despite the preference for hydrophobic amino acid residues

- 14 -

at 5 and 6, in some cases it may be desirable to have hydrophilic
amino acid residues at these positions. Specifically, amino acid
residue 5 may be a T, R or S, and amino acid residue 6 may be a T
or R. Likewise, while a hydrophilic amino acid residue is
preferred at position 9, in some instances a hydrophobic residue,
such as a P or A, may be desirable.

The present invention also contemplates SH3 domain-binding
peptides with a minimum length of 10, 11, 12, 13, 14, 15 or more
amino acids. Such peptides contain additional amino acid
residues flanking the core sequence of R-2-L-P-5-6-P (SEQ ID
NO:71) either at the C-terminal end, the N-terminal end or both.
Thus, for example, such peptides include those that further
comprise a C-terminal-flanking amino acid sequence of the formula
10, 10-11, 10-11-12, 10-11-12-13 (SEQ ID NO:12) or 10-11-12-13-14
(SEQ ID NO:13), in which each number represents any amino acid
residue except cysteine, such that the amino acid residue 10 is
bound to the amino acid residue 9 by a peptide bond. In that
case, specific embodiments include an amino acid residue 10 which
is T, R, L, S, D, P, A or N, preferably T or R, an amino acid
residue 11 which is R, P, A, Q, S or T, preferably R or P, an
amino acid residue 12 which is P, S, R or T, preferably P or S,
an amino acid residue 13 which is P, S, R, F, H or T, preferably
P or S, and an amino acid residue 14 which is S, R, G or T,
preferably, S or R.

Furthermore, peptides are also provided which further
comprise an N-terminal-flanking amino acid sequence of the
formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in
which each number represents any amino acid residue except
cysteine, such that 1' is bound to R by a peptide bond. In such
a case, specific embodiments are provided in which the amino acid
residue 1' is T, P, S, N, F, W, K, H, Q or G, preferably T or P,
wherein the amino acid residue 2' is S, T, G, P, R, Q, L, A or H,
preferably S or T, wherein the amino acid residue 3' is R, S, P,
G, A, V, Y or L, preferably S or T, and wherein the amino acid
residue 4' is R, S, V, T, G, L or F, preferably R or S.

- 15 -

In a particular embodiment, a peptide is disclosed having at least thirteen and up to forty-five amino acid residues, including an amino acid sequence of the formula, 3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned anywhere

5   along the peptide, in which each number represents an amino acid residue, such that 3', 2', 1', 2, 8, and 10 each represents any amino acid residue except cysteine, 5 and 6 each represents a hydrophobic amino acid residue, and 9 represents a hydrophilic amino acid residue except cysteine, each letter being the

10 standard one-letter symbol for the corresponding amino acid, said peptide exhibiting a binding affinity for the SH3 domain of Src. Preferred 13-mers include, but are not limited to, those having an amino acid residue 5 which is a P or M, an amino acid residue 1' which is T, P, S or N, an amino acid residue 2' which is S or

15 T, an amino acid residue 3' which is R or S, and an amino acid residue 10 which is T or R.   In all the SH3 domain-binding peptides described herein, the prohibition against the use of the hydrophilic amino acid residue cysteine (C) does not extend beyond the 7-mer "core" sequence and the additional amino acid

20 residues flanking the core up to a total (core + flanking) of about 20 amino acids. That is, the occasional use of a cysteine is not absolutely prohibited. What should be kept in mind is that the potential for the formation of intramolecular disulfide bonds, to form a cyclic structure, be minimized as much as

25 possible. Applicants have found that cyclized structures appear to be disfavored, at least with potential binding peptides of less than about 15 amino acid residues in length. The concern for the formation of cyclized structures comprising the core sequence diminishes with increasing size of the peptide.

30 Presumably, a large enough structure, though cyclic, may allow the critical core sequence to adopt a more or less linear conformation.

In particular, specific peptides are disclosed which exhibit binding affinities to SH3 domains. These include the peptides,

35 RSTPRPLPMLPTTR (SEQ ID NO. 62), RSTPRPLPPLPTTR (SEQ ID NO. 67),

- 16 -

GILAPPVPPRNTR (SEQ ID NO. 63), VLKRPLPIPPVTR (SEQ ID NO. 64),
GPHRRLPPTPATR (SEQ ID NO. 65), and ANPSPATRPLPTR (SEQ ID NO. 66).

Phage clones are also disclosed, along with the amino acid
sequences that are responsible for SH3 domain binding. These

5   phage clones are identified in Figure 5.

In other embodiments of the present invention, SH3 domain-
binding peptides are contemplated which have a total of 11, 13,
14, 18, 20, 22, 23, 25, 30, 36, 38 or 45 amino acid residues.

The peptides of the present invention, having been disclosed

10   herein, may be prepared by any number of practicable methods,
including but not limited to solution-phase synthesis, solid-
phase synthesis, protein expression by a transformed host,
cleavage from a naturally-derived, synthetic or semi-synthetic
polypeptide, or a combination of these techniques.

15   The SH3 binding peptides exhibit a wide range of biological
activity which includes the enhancement (or inhibition, depending
on the particular peptide or the nature of the peptide's target
molecule, in this case a protein bearing an SH3 domain) of the
natural function or biological activity of the peptide's target

20   molecule. For example, the interaction of the binding peptide of
the present invention could result in the modulation of the
oncogenic activity of the target molecule bearing the SH3 domain.
If the target molecule has, in turn, a natural binding partner or
ligand, the peptides of the present invention may also exhibit

25   antagonistic or agonistic activity in relation to the biological
activity of the natural binding partner.

Thus, it is an object of the present invention to provide a
method of activating Src or Src-related protein tyrosine kinases
by administering an effective amount of the SH3 domain-binding

30   peptides generally described herein. The intensity of the immune
response can thus be stimulated, for example, by the increased
production of certain lymphokines, such as TNF-alpha and
interleukin-1. As is generally known to those of ordinary skill
in the art, a more intense immune response may be in order in

certain conditions, such as in combating a particularly tenacious
infection, viral or otherwise, or a malignancy.

Furthermore, in a specific embodiment of the present
invention, a conjugate compound is contemplated which comprises
the peptide of the present invention and a second chemical
moiety. The second chemical moiety can be selected from a wide
variety of chemical compounds including the peptide itself.
Typically, however, the second chemical moiety is selected to be
other than the peptide of the present invention, including but
not limited to an amino acid, a peptide other than an SH3 binding
peptide of the present invention, a polypeptide or protein (i.e.,
the conjugate is a fusion protein), a nucleic acid, a nucleoside,
a glycosidic residue (i.e., any sugar or carbohydrate), a label
or image-enhancing agent (including metals, isotopes,
radioisotopes, chromophores, fluorophores (such as FITC, TRITC,
and the like), and enzyme substrates), a drug (including
synthetic, semisynthetic, and naturally-occurring compounds),
small molecules (e.g., biotin, hormones, factors) and the like.

The peptide of the present invention can be conjugated to
the second chemical moiety either directly (e.g., through
appropriate functional groups, such as an amine or carboxylic
acid group to form, for example, an amine, imine, amide, ester,
acyl or other carbon-carbon bond) or indirectly through the
intermediacy of a linker group (e.g., an aliphatic or aromatic
polyhydroxy, polyamine, polycarboxylic acid, polyolefin or
appropriate combinations thereof). Moreover, the term
"conjugate," as used herein, is also meant to encompass non-
covalent interactions, including but not limited to ionic,
affinity or other complexation interactions. Preferably, such
other non-covalent interactions provide definable, most
preferably, isolatable chemical conjugate species.

As described further herein, the peptides of the present
invention have been shown to localize within certain cellular
compartments which contain Src or Src-related prot ins.
Consequently, the above-described conjugate can be utilized as a

- 18 -

delivery system for introduction of a drug to cells, tissues or
organs that include SH3 domain-containing proteins.

It should also be pointed out that the present invention
seeks to provide a recombinant construct comprising a nucleic
5   acid or its complement that includes codons or nucleotide
sequences encoding a peptide having a region that binds to an SH3
domain, preferably the Src SH3 domain. The recombinant nucleic
acid may be a DNA or RNA polynucleotide.

In a specific embodiment, the present invention contemplates
10   a recombinant construct which is a transforming vector. Such
vectors include those well known to those of ordinary skill in
the art, which effect the transfer or expression of the
nucleotide sequence after introduction to a host, such as
recombinant plasmid, phage or yeast artificial chromosome. These
15   vectors may be closed circular loops or they may be linearized.
The vectors contemplated include those that exist
extrachromosomally after host transformation or transfection, as
well as those that integrate within or even displace portions of
the host chromosome. The vectors may be introduced to the cell
20   with the help of transfection aids or techniques well-known in
the art. For example, these aids or techniques may take the form
of electroporation, use of calcium chloride, calcium phosphate,
DEAE dextran, liposomes or polar lipid reagents known as
LIPOFECTIN or LIPOFECTAMINE. In addition, the present invention
25   contemplates the direct introduction of the desired nucleic acid
to the host cell, for instance, by injection.

Transformed host cells are also obtained by the methods of
the present invention which are capable of reproducing the
polynucleotide sequences of interest and/or expressing the
30   corresponding peptide products. A variety of hosts are
contemplated, including prokaryotic and eukaryotic hosts. In
particular, bacterial, viral, yeast, animal, and plant cells are
potentially transformable hosts. Thus, a method is disclosed to
obtain a transformed host cell that can produce, preferably
35   secrete, a peptide having a region that binds to an SH3 domain

- 19 -

comprising (a) providing an expression vector, preferably a
secretory expression vector, comprising a nucleotide sequence
encoding at least one copy of a peptide having a region that
binds to an SH3 domain; and (b) introducing the vector to a
5   competent host cell.

The peptides, thus produced, may then be introduced to
cells, tissues, organs, or administered to the subject for the
purpose of modulating the biochemical activity of the SH3 domain-
containing proteins present therein. Accordingly, in specific
10   embodiments of the present invention, compositions are provided
which comprise an SH3 domain-binding peptide, including a core
sequence and flanking sequences, and a suitable carrier.

The compositions contemplated by the present invention may
also include other components, from those that facilitate the
15   introduction or administration of the compositions to those that
have their own innate activity, such as a prophylactic, a
diagnostic or a therapeutic action. Such innate activity may be
distinct from that of the peptides of the present invention or be
complementary thereto. In any event, the compositions of the
20   present invention include those that are suitable for
administration into mammals, including humans. Preferably, the
compositions (including necessarily the carrier) of the present
invention are sterile, though others may need only be
cosmetically, agriculturally or pharmaceutically acceptable.
25   Still other compositions may be adapted for veterinary use.

The compositions, including the drug delivery systems
described herein, are contemplated to be administered in a
variety of ways, such as parenterally, orally, enterally,
topically or by inhalation. The compositions may also be
30   adminstered intranasally, opthalmically or intravaginally.
Furthermore, the compositions of the invention can take several
forms, such as solids, gels, liquids, aerosols or patches.

In another embodiment of the present invention a method is
provided of identifying a peptide having a region that binds to
35   an SH3 domain comprising: (a) providing an immobilized target

- 20 -

protein comprising an SH3 domain; (b) incubating the immobilized
target protein with an aliquot taken from a phage-displayed
random peptide library, which library includes peptides having a
random sequence of ≥8 amino acid residues; (c) washing unbound
5   phage from the immobilized target protein; (d) recovering the
phage bound to the immobilized target protein; and (e)
determining the relevant nucleotide sequence of said binding
phage nucleic acid and deducing the primary sequence
corresponding to the SH3 domain-binding peptide. Preferably, the
10   method further comprises amplifying the titer of the recovered
phage and repeating the steps of incubation, washing and recovery
to provide SH3 domain-binding peptide-enriched phage.

        Any other mode by which the peptide library, random or
otherwise, can be "displayed" can be utilized in the present
15   invention, however. Moreover, the present applicants believe
that longer *random* peptide sequences (e.g., >6 amino acid
residues, preferably >10, and most preferably, >12) provide not
only much greater diversity but also a richer degree of secondary
structure conducive to binding activity. If the random region of
20   the peptide is less than or equal to an 8-mer, it should
preferably not be cyclized.

5.2. Preparation of Random Peptide Libraries

        The preparation and characterization of the preferred
phage-displayed random peptide libraries have been described
25   elsewhere. See, for example, Kay, B.K. et al. in Gene (1992)
128:59-65, for a description of the preparation of the phage-
displayed random peptide library known as TSAR-9, more below. In
particular, by cloning degenerate oligonucleotides of fixed
length into bacteriophage vectors, recombinant libraries of
30   random peptides can be generated which are expressed at the
amino-terminus of the pIII protein on the surface of M13 viral
particles. (There are 3-5 copies of the pIII-fusion on the
surface of each particle.) Phage display offers several
conveniences: first, the expressed peptides are on the surface

                                - 21 -

of the viral particles and accessible for interactions; second, the recombinant viral particles are stable (i.e., can be frozen, exposed to pH extremes); third, the viruses can be amplified; and fourth, each viral particle contains the DNA encoding the

5   recombinant genome. Consequently, these libraries can be screened by isolating viral particles that bind to targets. The isolates can be grown up overnight, and the displayed peptide sequence responsible for binding can be deduced by DNA sequencing.

10   These libraries have approximately >10^8 different recombinants, and nucleotide sequencing of the inserts suggests that the expressed peptides are indeed random in amino acid sequence. These libraries are referred to herein as TSAR libraries, where TSAR stands for Totally Synthetic Affinity

15   Reagents. The preparation of the TSAR libraries are described further below.

5.3. SH3 Binding Clones And Their Characteristics

Accordingly, peptides have been isolated from an unconstrained random peptide library which exhibit a binding

20   affinity for SH3 domains. Furthermore, the binding affinities exhibited by the disclosed peptides differ in their selectivities with certain peptides showing comparable binding affinities for SH3 domains derived from different proteins, while others manifest greater affinities for specific SH3 domains.

25   The amino acid sequence of various peptides isolated by the present method are listed in Figure 5. As can be seen from this list, certain groups of SH3 domain binding peptides are isolated from three separate random peptide libraries, each based on a different type of random peptide insert, all displayed at the

30   amino-terminus of the pIII protein on the surface of M13 viral particles. Ten clones were isolated from the R8C library, seven from the TSAR-12 library, and seven from the TSAR-9 library. The sequences are presented to highlight the particular amino acid residues believed to bind directly to the SH3 domain, as well as

- 22 -

to point out the remaining amino acid resiudes of the random
insert and the viral flanking sequences and complementary site
amino acid residues common to each group of clones. The
frequency with which each particular clone is found in each
5   library is also indicated in Figure 5. Thus, clones T12.SRC3.1
and T12.SRC3.2 are by far the most abundant clones found among
the three libraries.

Interestingly, all the binding peptides are found to have
the proline-rich amino acid residue motif, which is apparently
10   responsible for binding, the motif being located predominantly at
the C-terminal end of the insert, although each clone also
contains an insert at the N-terminal end. The significance of
this observation is not presently understood, although this
finding may indicate the possible importance of the C-terminal
15   viral flanking sequences in SH3 domain binding.

Indeed, a synthetic peptide bearing only the core consensus
sequence RPLPPLP (SEQ ID NO:9) was less effective in binding to
target SH3 domains than synthetic peptides that also included
additional amino acid residues flanking the core sequences.
20   Thus, 13-mers and 14-mers having the sequences RSTPRPLPMLPTTR
(SEQ ID NO:62), RSTPRPLPPLPTTR (SEQ ID NO:67), GILAPPVPPRNTR (SEQ
ID NO:63), GPHRRLPPTPATR (SEQ ID NO:65), and VLKRPLPIPPVTR (SEQ
ID NO:64) have been prepared and shown to bind to SH3 domains,
such as those of Src and Yes, much more avidly than the 7-mer,
25   RPLPPLP (SEQ ID NO:9). The 13-mer ANPSPATRPLPTR (SEQ ID NO:66)
has been shown to have binding affinities comparable to the core
consensus sequence. In each case, the 13-mers comprise a 7-mer
"core" sequence plus additional amino acid residues flanking
same, some of which additional amino acid residues are
30   contributed by the viral flanking sequences.

Thus, in one embodiment of the present invention, a 7-mer
core includes a consensus motif of the formula RXLP$\phi\phi$P (SEQ ID
NO:71), wherein R is arginine, L is leucine, P is proline, X
represents any amino acid except cysteine and ϕ represents a
35   hydrophobic amino acid residue. By "hydrophobic amino acid

- 23 -

residue," the applicants mean to include F, Y, W, V, A, I, L, P
or M, each letter representing the standard one-letter
designation for the corresponding amino acid residue.

5   Furthermore, a preferred 9-mer peptide comprising two
additional amino acids on the C-terminal end of the core sequence
is envisioned having a consensus motif of the formula RXLPφφPXψ.
In this preferred 9-mer consensus motif, the symbol ψ represents
a hydrophilic amino acid residue, except cysteine. By
"hydrophilic amino acid residue," the applicants mean to include

10   K, R, H, D, E, N, Q, T, S or C, and the other symbols are as
defined above. For the purposes of the present invention, a
glycine residue (G) may be considered either a hydrophobic or a
hydrophilic amino acid residue. The one-letter symbols B and Z,
which stand for N or D and Q or E, respectively, are considered

15   hydrophilic amino acid residues.

Particular 13-mer peptides of the present invention include
those listed, below. It is noted, however, that not all the
following 13-mer peptides correlate strictly to or comply with
the preferred 9-mer consensus motif, described above. Those

20   peptides that *do not* comply (indicated in italics, with the non-
complying amino acid residues underscored) can, thus, be
described as "resembling" those that do comply (indicated in
normal type) with the preferred 9-mer consensus motif:

PGFRELPPLPPSR (SEQ ID NO:72),   AQSRPLPIPPETR (SEQ ID NO:73),

25   VLKRPLPIPPVTR (SEQ ID NO:64),   *PPNSPLPPLPTHL* (SEQ ID NO:74),

TGRGPLPPLPNDS (SEQ ID NO:75),   *YSTRPVPPITRPS* (SEQ ID NO:76),

SHKSRLPPLPTRP (SEQ ID NO:77),   *YRFRALPSPPSAS* (SEQ ID NO:78),

GPHRRLPPTPATR (SEQ ID NO:65),   *LAQRQLPPTPGRD* (SEQ ID NO:79),

ALQRRLPRTPPPA (SEQ ID NO:80),   *PATRPLPTRPSRT* (SEQ ID NO:81),

30   *YSTRPLPSRPSRT* (SEQ ID NO:82),   *XPGRILLLPSEPR* (SEQ ID NO:83),

SGGILAPPVPPRN (SEQ ID NO:84),   RSTRPLPILPRTT (SEQ ID NO:85),

STPRPLPMLPTTR (SEQ ID NO:86),   STNRPLPMIPTTR (SEQ ID NO:87),

RSTRPLPSLPITT (SEQ ID NO:88),   *STSRPLPSLPTTR* (SEQ ID NO:89),

RSTRSLPPLPPTT (SEQ ID NO:90),   RSTRQLPIPPTTT (SEQ ID NO:91),

35   STPRPLPLIPTTP (SEQ ID NO:92),   *RSTRPLPPTPLTT* (SEQ ID NO:93), and

- 24 -

RSTRPQPPPPITT (SEQ ID NO:94). Accordingly, other peptides not specifically disclosed, which either comply with or "resemble" the preferred 9-mer consensus motif, can be readily envisioned by those of ordinary skill in the art and are considered to be

5 equivalent to those that are specifically disclosed above. In particular, non-compliance at positions 1 (S, G, and I, in place of R, are tolerated), 3 (V, A, and Q, in place of L, are tolerated), 4 (L, in place of P, is tolerated), 5 (hydrophilic amino acid residues, S, R, and T, are tolerated in place of a

10 hydrophobic amino acid residue), 6 (hydrophilic amino acid residues, R and T, are tolerated in place of a hydrophobic amino acid residue), 7 (T, and S, in place of P, are tolerated), and 9 (P and A are tolerated in place of a hydrophilic amino acid residue) have been observed.

15   **5.3.1.   Binding Specificities**

It has been discovered that certain of the binding peptides disclosed have a greater relative binding affinity for one SH3 domain over another. Referring now to Figure 8, the relative binding affinities of the various peptides described

20 above toward different SH3 domain targets are graphically presented. As one can see, the relative binding affinities of the respective peptides can differ by orders of magnitude. Thus, as shown in Figure 8, the peptide GPHRRLPPTPATR (SEQ ID NO:65), having the relevant sequence of the phage clone identified as

25 T12.SRC3.3, is specific to Src family SH3 domains, including, but not limited to, Src, Yes, Lck, Hck, Fgr, Fyn, and Lyn. This SH3 binding peptide has little affinity for SH3 domains derived from PLCγ or Grb2. On the other hand, the peptide GILAPPVPPRNTR (SEQ ID NO:63), corresponding to the relevant sequence of the phage

30 clone T12.SRC3.1, which is one of the most abundant binding clones found by the present method, binds generically to a broad range of SH3 domains, including Src, PLCγ, and Grb2.

On an intermediate level, the present invention has also uncovered a peptide, VLKRPLPIPPVTR (SEQ ID NO:64), corresponding

to the relevant sequence of the phage clone T12.SRC3.6, which is
Src preferential; that is, this peptide exhibits strong binding
affinities for members of the Src family, some binding affinities
for Grb2 proteins, but little binding affinities for PLCγ

5   domains. The peptide ANPSPATRPLPTR (SEQ ID NO:66), corresponding
to the relevant sequence of the phage clone T12.SRC3.2, also
exhibits Src family specificity similar to GPHRRLPPTPATR (SEQ ID
NO:65). The peptides RSTPRPLPMLPTTR (R8C.YES3.5; SEQ ID NO:62)
and RSTPRPLPPLPTTR (representative consensus motif; SEQ ID NO:67)

10   are highly specific for SH3 domain of Src, Yes, and other Src-
related proteins.

5.4. Further Discussion of Binding Experiments

At the outset it is apparent that the binding affinity
of certain peptides to the SH3 domain of Src and Src-related

15   proteins is governed by more than just the presence of the
preferred core consensus sequences, RPLPPLP (SEQ ID NO:9) or
RPLPMLP (SEQ ID NO:95; i.e., RPLP(P/M)LP, SEQ ID NO:96). Thus,
while the synthetic peptides RSTPRPLPMLPTTR (R8C.YES3.5; SEQ ID
NO:62) and RSTPRPLPPLPTTR (consensus; (SEQ ID NO:67) exhibit a

20   strong specific binding affinity for Src SH3, the other synthetic
peptides tested also exhibited an avid binding affinity to SH3
domains relative to the 7-mer, RPLPPLP (SEQ ID NO:9). These
other peptides, GILAPPVPPRNTR (SEQ ID NO:63), VLKRPLPIPPVTR (SEQ
ID NO:64), GPHRRLPPTPATR (SEQ ID NO:65), and ANPSPATRPLPTR (SEQ

25   ID NO:66), sport core sequences and flanking sequences that do
not closely adhere to the preferred core consensus sequences.
Thus, these results suggest that binding affinity to SH3 domains
is governed to a large extent by the nature of the amino acid
residues flanking the core 7-mer sequence.

30   The binding characteristics of Src SH3-selected peptides was
determined using synthetic biotinylated peptides corresponding to
the sequences displayed by Src SH3-selected phage. These
biotinylated peptides were assayed for direct binding to
immobilized Src SH3-GST. Each of the five library-derived

peptides tested were found to bind to Src SH3-GST and Yes SH3-GST
over background (Figure 8). Furthermore, a strong correlation
was observed between the similarity of a given peptide to the
preferred core consensus sequence RPLP(P/M)LP and the peptide's
5   affinity for Src SH3-GST. The core sequence of the clone
T12.SRC3.1 (GILAPPVPPRNTR; SEQ ID NO:63) appears to provide more
generic SH3 domain-binding characteristics.

Experiments comparing the relative binding of various phage
clones to SH3 domains taken from a variety of proteins
10   demonstrated the preference of these clones for Src and Src-
related SH3 domains over SH3 domains taken from other proteins.

It was further found that while the 7-mer having the
consensus sequence RPLPPLP (SEQ ID NO:9) bound to Src SH3-GST
only weakly, peptides comprising the consensus sequence flanked
15   by residues encoded by one of the Src SH3-selected clones
(R8C.YES3.5), RSTP (SEQ ID NO:97) at the N-terminal end and TTR
at the C-terminal end, bound significantly better than any of the
peptides tested (Figure 7). Thus, as stated previously,
sequences that flank the RPLP(P/M)LP (SEQ ID NO:96) core appear
20   to be important contributors to SH3 binding. It is further
surmised that a peptide having or resembling the sequence
RSTPAPPVPPRTTR (SEQ ID NO:98) should exhibit strong but generic
binding to a variety of SH3 domains.

Similarly, it is observed that most of the Src SH3-binding
25   motifs are located near the carboxy-terminus of the random
peptides, adjacent to sequences which are fixed in every clone
(Figure 5). The exceptional clones tend to possess sequences
that resemble motifs that include fixed flanking sequences. This
clustering contrasts with previous results, in which binding
30   motifs are distributed throughout the random peptide. Kay, B.K.,
et al., in _Gene_ (1993) 128:59-65.

The binding of the library-derived Src SH3-binding peptides
was compared to that of peptides corresponding to proline-rich
regions of natural proteins. Peptides corresponding to SH3-
35   binding regions in human PI-3' Kinase (KISPPTPKPRPPRPLPV; SEQ ID

- 27 -

NO:69) and human SOS1.20 (GTVEPVPPPVPPRRRPESA; SEQ ID NO:68), as
well as a proline-rich region of the cytoskeletal protein
vinculin (LAPPKPPLPEGEV; SEQ ID NO:70), bound Src SH3 less well
than the library-derived peptides (Figure 7).

5   As mentioned above, the relative specificity of binding was
explored. Thus, the relative binding of Src SH3-selected
peptides to equal amounts of GST fusions to SH3 domains from
different proteins was determined (Figure 8). While all of the
library-derived peptides bound the Src and Yes SH3 domains almost
10   equally well, none of the peptides (with the exception of peptide
T12.SRC3.1, the most divergent peptide tested) bound the SH3
domains of Grb2, Crk, Abl or PLCγ1 appreciably. Thus, the
library-derived peptides, in contrast with a peptide derived from
SOS1, exhibit SH3 binding that is relatively specific for Src-
15   family members.

        Next, it was determined whether the binding to the Src SH3
domain was qualitatively like the interactions of the SH3 domain
and natural proteins found in cell lysates. Thus, radiolabeled
proteins were prepared from NIH 3T3 cell lysates and
20   chromatographed over Src SH3-GST immobilized on glutathione
linked Sepharose. SDS-PAGE shows that a number of proteins can
be affinity purified in this manner. The synthesized peptides
bind quite well to the Src SH3 domain, as they can compete the
binding of radiolabeled proteins from cell lysates to immobilized
25   Src-GST, with an IC_{50} of 1-10 mM (Figure 9). In conclusion, the
peptides can efficiently block the interaction of cellular
proteins with Src SH3 *in vitro*.

        Moreover, *Xenopus laevis* oocytes injected with mRNA encoding
constitutively active Src undergo progesterone-induced maturation
30   at an accelerated rate relative to oocytes injected with water or
c-Src mRNA. Unger, T.F. and Steele, R.E. in Mol. Cell.Biol.
(1992) 12:5485-5498. To explore the ability of the library-
derived Src SH3-binding peptides to exert a biochemical effect *in
vivo*, the influence of the peptides on the maturation of *Xenopus
35   laevis* oocytes was examined. Hence, stage VI oocytes were

- 28 -

injected with peptide, exposed to progesterone, and scored for
germinal vesicle breakdown. Figure 10 shows that the rate of
maturation was accelerated by approximately one hour when oocytes
were injected with the SH3-binding peptide consisting of RPLPPLP

5   (SEQ ID NO:9) flanked by residues from clone T12.SRC3.6
(VLKRPLPIPPVTR; SEQ ID NO:64), but not with water or a peptide
corresponding to a proline-rich segment of vinculin
(LAPPKPPLPEGEV; SEQ ID NO:70) as controls. The magnitude of this
effect is roughly equivalent to that seen with injection of mRNA

10   encoding constituitively active Src. See, e.g., Figure 3B in
Unger, T.F. and Steele, R.E., *supra*. This result suggests that
the library-derived Src SH3-binding peptide is effectively
relieving an inhibitory effect of the Src SH3 domain upon Src PTK
activity. This model is consistent with a number of studies

15   which have demonstrated an inhibitory effect of the Src SH3
domain upon Src kinase and transforming activity. See, e.g.,
Okada, M., et al., *supra*; Murphy, S.M., et al., *supra*; and
Superti-Furga, G., et al., *supra*.

       5.5. Diagnostic And Therapeutic Agents
20   **Based On SH3 Binding Peptides and Additional Methods**
                of Their Use
               As already indicated above, the present invention also
seeks to provide diagnostic, prophylactic, and therapeutic agents
based on the SH3 binding peptides described herein.

25   In one embodiment, diagnostic agents are provided,
preferably in the form of kits, comprising an SH3 domain-binding
peptide and a detectable label conjugated to said peptide
directly, indirectly or by complexation, said peptide comprising:
(i) a core sequence motif of the formula RXLP$\phi\phi$P (SEQ ID NO:71),

30   wherein X represents any amino acid except cysteine and ϕ
represents a hydrophobic amino acid residue, including F, Y, W,
V, A, I, L, P, M or G, each letter representing the standard one-
letter designation for the corresponding amino acid residue; and

- 29 -

(ii) two or more additional amino acid residues flanking said
core sequence at its C-terminal end, N-terminal end or both.

The diagnostic agents of the present invention can be used
to detect the presence of SH3 domains of a generic or specific

5   type in cells, tissues or organs either *in vitro* or *in vivo*. For
in vivo applications, the diagnostic agent is preferably mixed
with a pharmaceutically acceptable carrier for administration,
either enterally, parenterally or by some other route dictated by
the needs of the particular application.

10   In a particular embodiment, for example, an assay based on a
fusion product is contemplated which comprises a Src SH3 domain-
binding peptide of the invention and a substrate for deregulated
or "activated" Src. For instance, a muscle biopsy, taken from a
subject suspected of being infected by the Rous sarcoma virus,

15   can be treated with an effective amount of the fusion product.
By subsequent analysis of the degree of conversion of the
substrate, one can potentially detect infection by the Rous
sarcoma virus in the subject, particularly mammals, especially
chickens. The presence of the retrovirus, which causes the

20   expression of deregulated or "activated" Src, may thus be
indicated by unusually high levels of Src as revealed by large
amounts of the converted substrate. See, for example, Paxton,
W.G. et al., in <u>Biochem. Biophys. Res. Commun.</u> (1994) 200(1):260-
267 (detection of phosphorylated tyrosine and serine residues of

25   angiotensin II AT1 receptor, a substrate of Src family tyrosine
kinases); another suitable substrate may be the protein p68
(Fumagalli, S. et al., in <u>Nature</u> (1994) 368(6474):871-874;
Taylor, S.J. and Shalloway, D., in <u>Ibid.</u> at 867-871.

Alternatively, the enzyme can be isolated by selective

30   binding to a form of the SH3 domain-binding peptides of the
present invention (e.g., biotin-peptide conjugate). After
isolation of the protein-peptide conjugate complex (e.g., on a
column comprising streptavidin), the activity of the enzyme can
then be assayed by conventional methods to determine its level of

35   protein kinase activity which can be taken as an indication of

- 30 -

the presence of the deregulated or "activated" form of the
enzyme. An assay for Src kinase has been described by Klinz and
Maness, in Neuroprotocols (a companion to Neuroscience)(1992)
1(3):224-231.

5   Moreover, the diagnostic agents of the invention can also
serve as imaging agents of cells, tissues or organs, especially
those that contain proteins with an SH3 domain. For example,
neural cells (e.g., neurons, other areas of the brain),
osteoclasts, osteoblasts, platelets, immune cells, and other
10   dividing cells are known to express or contain proteins with SH3
domains. Thus, an image can be taken of portions of the body to
serve as a baseline for subsequent images to detect physiologic
or biochemical changes in the subject's body. For instance,
changes in the condition of cellular levels of Src or a
15   transformation of the cellular Src to an "activated" form may be
detected using the diagnostic or imaging agents of the present
invention.

           Accordingly, it has been demonstrated that an SH3-binding
peptide tagged with a fluorescence emitter can provide an image
20   of the cytoskeleton. The images are presented in Figure 11. As
can be seen from Figure 11, panels A, B, and C show the
fluorescence image that is obtained on treating NIH 3T3
fibroblasts with SH3 domain-binding peptides modified to include
a fluorescent tag. In sharp contrast, panel D shows only a dark
25   image that is produced when the cells are treated with a proline-
rich segment of vinculin as a control.

           In another embodiment, an SH3 domain-binding peptide-
horseradish immunoperoxidase complex or related
immunohistochemical agent could be used to detect and quantitate
30   specific receptor molecules in tissues, serum or body fluids. In
particular, the present invention provides useful diagnostic
reagents for use in immunoassays, Southern or Northern
hybridization, and in situ assays. Accordingly, the diagnostic
agents described herein may be suitable for use in vitro or in
35   vivo.

In addition, the diagnostic or imaging agent of the present
invention is not limited by the nature of the detectable label.
Hence, the diagnostic agent may contain one or more such labels
including, but not limited to, radioisotope, fluorescent tags,
5   paramagnetic substances, heavy metals, or other image-enhancing
agents. Those of ordinary skill in the art would be familiar
with the range of label and methods to incorporate or conjugate
them into the SH3 domain-binding peptide to form diagnostic
agents.

10   In yet a further embodiment, pharmaceutical compositions are
provided comprising an SH3 domain-binding peptide and a
pharmaceutically acceptable carrier. In a specific embodiment of
the invention, the pharmaceutical composition is useful for the
modulation of the activity of SH3 domain-containing proteins. By
15   "modulation" is meant either inhibition or enhancement of the
activity of the protein target.   Accordingly, a pharmaceutical
composition is disclosed comprising an SH3 domain-binding peptide
and a pharmaceutically acceptable carrier, said peptide
comprising: (i) a 9-mer sequence motif of the formula RXLP$\phi\phi$PXψ
20   (SEQ ID NO:10), wherein X represents any amino acid except
cysteine, ϕ represents a hydrophobic amino acid residue, and
wherein ψ is a hydrophilic amino acid residue except cysteine,
each letter representing the standard one-letter designation for
the corresponding amino acid residue; and, optionally, (ii)
25   additional amino acid residues flanking the 9-mer sequence at its
C-terminal end, N-terminal end or both, up to a total of 45 amino
acid residues, including said 9-mer sequence. Preferably, the
peptide comprises at least one, more preferably at least two, and
most preferably at least three additional amino acids flanking
30   the 9-mer sequence.

As stated above, the therapeutic or diagnostic agents of the
invention may also contain appropriate pharmaceutically
acceptable carriers, diluents and adjuvants. Such pharmaceutical
carriers can be sterile liquids, such as water and oils including
35   those of petroleum, animal, vegetable or synthetic origin, such

- 32 -

as peanut oil, soybean oil, mineral oil, sesame oil and the like.
Water is a preferred carrier when the pharmaceutical composition
is administered intravenously. Saline solutions and aqueous
dextrose and glycerol solutions can also be employed as liquid

5   carriers, particularly for injectable solutions. Suitable
pharmaceutical excipients include starch, glucose, lactose,
sucrose, gelatin, malt, rice, flour, chalk, silica gel, magnesium
carbonate, magnesium stearate, sodium stearate, glycerol
monostearate, talc, sodium chloride, dried skim milk, glycerol,

10   propylene, glycol, water, ethanol and the like. These
compositions can take the form of solutions, suspensions,
tablets, pills, capsules, powders, sustained-release formulations
and the like. Suitable pharmaceutical carriers are described in
"Remington's Pharmaceutical Sciences" by E.W. Martin.

15   Such compositions will contain an effective therapeutic
amount of the active compound together with a suitable amount of
carrier so as to provide the form for proper administration to
the subject. While intravenous injection is a very effective
form of administration, other modes can be employed, including

20   but not limited to intramuscular, intraperitoneal, and
subcutaneous injection, and oral, nasal, enteral, and parenteral
administration.
           The therapeutic agents and diagnostic agents of the instant
invention are used for the treatment and/or diagnosis of animals,

25   and more preferably, mammals including humans, as well as dogs,
cats, horses, cows, pigs, guinea pigs, mice and rats.
Accordingly, other methods contemplated in the present invention,
include, but are not limited to, a method of modulating, i.e.,
inhibiting or enhancing, bone resorption in a mammal (see, e.g.,

30   Hall, T.J., in <u>Biochem. Biophys. Res. Commun.</u> (1994) 199(3):1237-
44), a method of disrupting protein tyrosine kinase-mediated
signal transduction pathways or a method of regulating the
processing, trafficking or translation of RNA in a cell by
introducing or administering an effective amount of an SH3

- 33 -

domain-binding peptide of the present invention (see, e.g.,
Taylor, S.J. and Shalloway, D., *supra*).

5

10

15

20

25

30

35

The diagnostic or therapeutic agents of the present
invention can be modified by attachment to soluble macromolecules
such as proteins, polysaccharides, or synthetic polymers. For
example, the peptide could be coupled to styrene-maleic acid
copolymers (see, e.g., Matsumura and Maeda, Cancer Res. (1986)
46:6387), methacrylamide copolymers (Kopececk and Duncan, J.
Controlled Release (1987) 6:315), or polyethylene glycol (PEG)
(e.g., Hershfield and Buckley, N. Engl. J. Med. (1987) 316:589;
Ho et al., Drug Metab. Dispos. (1986) 14:349; Chua et al., Ann.
Intern. Med. (1988) 109:114).   The agents, if desired, are
further targeted by attachment to an antibody, especially a
monoclonal antibody. Such antibodies include but are not limited
to chimeric, single chain, Fab fragments, and Fab expression
libraries. In one embodiment the agent is coupled to the
macromolecule via a degradable linkage so that it will be
released *in vivo* in its active form.

In another embodiment, the therapeutic or diagnostic agent
may be delivered in a vesicle, in particular a liposome. See,
Langer, Science (1990) 249:1527-1533; Treat et al., in Liposomes
in the Therapy of Infectious Disease and Cancer, Lopez-Berestein
and Fidler (eds.), Liss, New York (1989) pp. 353-365; Lopez-
Berestein, ibid., pp. 317-327.

In yet another embodiment, the therapeutic or *in vivo*
diagnostic agent can be delivered in a controlled release system.
In one embodiment, a pump may be used (see Langer, *supra*; Sefton,
CRC Crit. Ref. Biomed. Eng. (1987) 14:201; Buchwald et al.,
Surgery (1980) 88:507; Saudek et al., N. Engl. J. Med. (1989)
321:574). In another embodiment, polymeric materials may be used
(see Medical Applications of Controlled Release, Langer and Wise
(eds.), CRC Pres., Boca Raton, Florida, 1974; Controlled Drug
Bioavailability, Drug Product Design and Performance, Smolen and
Ball (eds.) Wiley, New York 1984; Raner and Peppas, J. Macromol.
Sci. Rev. Macromol. Chem. (1983) 23:61; see, also, Levy et al.,

- 34 -

Science (1985) 228:190; During et al., <u>Ann. Neurol.</u> (1989)
25:351; Howard et al., <u>J. Neurosurg.</u> (1989) 71:105). In a
preferred embodiment, a controlled release system may be placed
next to the therapeutic target, thus requiring only a fraction of

5   the systemic dose (see, e.g., Goodson, in <u>Medical Applications of
Controlled Release</u>, *supra*, (1984) 2:115-138). It will be
recognized by one of ordinary skill in the art that a particular
advantage of the invention is that a peptide will not be subject
to the problems of denaturation and aggregation associated with

10   proteins held in the warm, most environment of a body in a
controlled release system.

Other controlled release systems are discussed in the review
by Langer, in <u>Science</u> (1990) 249:1527-1533.

6.   **EXAMPLES**

6.1. Preparation of the TSAR-9 Library

6.1.1.   Synthesis and Assembly of Oligonucleotides

5   Figure 1 shows the formula of the
oligonucleotides and the assembly scheme used in construction of
the TSAR-9 library. The oligonucleotides were synthesized with
an applied Biosystems 380a synthesizer (Foster City, CA), and the
full-length oligonucleotides were purified by HPLC.

10   Five micrograms of each of the pair of oligonucleotides were
mixed together in buffer (10 mM Tris-HCl, pH 8.3, 15 mM KCl,
0.001% gelatin, 1.5 mM magnesium chloride), with 0.1 % Triton X-
100, 2 mM dNTP's, and 20 units of Taq DNA polymerase. The
assembly reaction mixtures were incubated at 72 °C for 30 seconds
15   and then 30 °C for 30 seconds; this cycle was repeated 60 times.
It should be noted that the assembly reaction is not PCR, since a
denaturation step was not used. Fill-in reactions were carried
out in a thermal cycling, device (Ericomp, LaJolla, CA) with the
following protocol: 30 seconds at 72 °C, 30 seconds at 30 °C,
20   repeated for 60 cycles. The lower temperature allows for
annealing of the six base complementary region between the two
sets of the oligonucleotide pairs. The reaction products were
phenol/chloroform extracted and ethanol precipitated. Greater
than 90% of the nucleotides were found to have been converted to
25   double stranded synthetic oligonucleotides.

         After resuspension in 300 μL of buffer containing 10 mM
Tris-HCl, pH 7.5, 1 mM EDTA (TE buffer), the ends of the
oligonucleotide fragments were cleaved with Xba I and Xho I (New
England BioLabs, Beverly, MA) according to the supplier's
30   recommendations. The fragments were purified by 4% agarose gel
electrophoresis. The band of correct size was removed and
electroeluted, concentrated by ethanol precipitation and
resuspended in 100 μL TE buffer. Approximately 5% of the
assembled oligonucleotides can be expected to have internal Xho I
35   or Xba I sites; however, only the full-length molecules were used

in the ligation step of the assembly scheme. The concentration
of the synthetic oligonucleotide fragments was estimated by
comparing the intensity on an ethidium bromide stained gel run
along with appropriate quantitated markers. All DNA
5   manipulations not described in detail were performed according to
Maniatis, *supra*.

To demonstrate that the assembled enzyme digested
oligonucleotides could be ligated, the synthesized DNA fragments
were examined for their ability to self-ligate. The digested
10   fragments were incubated overnight at 18 °C in ligation buffer
with T4 DNA ligase. When the ligation products were examined by
agarose gel electrophoresis, a concatamer of bands was visible
upon ethidium bromide staining. As many as five different unit
length concatamer bands (i.e., dimer, trimer, tetramer, pentamer,
15   hexamer) were evident, suggesting that the synthesized DNA
fragments were efficient substrates for ligation.

6.1.2.   Construction of Vectors

The construction of the M13 derived phage vectors
useful for expressing a TSAR library has been recently described
20   (Fowlkes, D. et al. <u>BioTech.</u> (1992) 13:422-427). To express the
TSAR-9 library, an M13 derived vector, m663, was constructed as
described in Fowlkes. The m663 vector contains the pIII gene
having a *c-myc*-epitope, i.e., as a stuffer fragment, introduced
at the mature N-terminal end, flanked by *Xho* I and *Xba* I
25   restriction sites (see also, Figure I of Fowlkes).

6.1.3.   Expression of the TSAR-9 Library

The synthesized oligonucleotides were then ligated
to *Xho* I and *Xba* I double-digested m663 RF DNA containing, the
pIII gene (Fowlkes) by incubation with ligase overnight at 12 °C.
30   More particularly, 50 ng of vector DNA and 5 ng of the digested
synthesized DNA and was mixed together in 50 μL ligation buffer
(50 mM Tris, pH 8.0, 10 mM $MgCl_2$, 20 mM DTT, 0.1 mM ATP) with T4
DNA ligase. After overnight ligation at 12°C, the DNA was

- 37 -

concentrated by ethanol precipitation and washed with 70%
ethanol. The ligated DNA was then introduced into *E. coli*
(DH5αF'; GIBCO BRL, Gaithersburg, MD) by electroporation.

      A small aliquot of the electroporated cells was plated and

5   the number of plaques counted to determine that 10^8 recombinants
were generated. The library of *E. coli* cells containing
recombinant vectors was plated at a high density (~400,000 per
150 mM petri plate) for a single amplification of the recombinant
phage. After 8 hr, the recombinant bacteriophage were recovered

10   by washing each plate for 18 hr with SMG buffer (100 mM NaCl, 10
mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.05% gelatin) and after the
addition of glycerol to 50% were frozen at -80 °C. The TSAR-9
library thus formed had a working titer of ~2 x 10^{11} pfu/ml.

6.2. Preparation of TSAR-12 Library

15   Figure 2 shows the formula for the synthetic
oligonucleotides and the assembly scheme used in the construction
of the TSAR-12 library. As shown in Figure 2, the TSAR-12
library was prepared substantially the same as the TSAR-9 library
described in Section 6.1 above with the following exceptions:

20   (1) each of the variant non-predicted oligonucleotide sequences,
i.e., NNB, was 30 nucleotides in length, rather than 54
nucleotides; (2) the restriction sites included at the 5' termini
of the variant, non-predicted sequences were *Sal* I and *Spe* 1,
rather than *Xho* I and *Xba* I; and (3) the invariant sequence at

25   the 3' termini to aid annealing of the two strands was GCGGTG and
CGCCAC rather than CCAGGT and GGTCCA (5' to 3').

      After synthesis including numerous rounds of annealing and
chain extension in the presence of dNTP's and <u>Taq</u> DNA polymerase,
and purification as described above in Section 6.1.1, the

30   synthetic double stranded, oligonucleotide fragments were
digested with *Sal* I and *Spe* I restriction enzymes and ligated
with T4 DNA ligase to the nucleotide sequence encoding the M13
pIII gene contained in the m663 vector to yield a library of
TSAR-expression vectors as described in Sections 6.1.2 and 6.1.3.

- 38 -

The ligated DNA was then introduced into *E. coli* (DH5αF'; GIBCO
BRL, Gaithersburg, MD by electroporation. The library of *E. coli*
cells were plated at high density (~400,000 per 150 mm petri
plate) for amplification of the recombinant phage. After about 8
5   hr, the recombinant bacteriophage were recovered by washing, for
18 hr with SMG buffer and after the addition of glycerol to 50%
were frozen at -80 °C.

The TSAR-12 library thus formed had a working titer of ~2 x
10^{11} pfu/mL.

10   **6.3. Characterization of the TSAR-9 and -12 Libraries**

The inserted synthetic oligonucleotides for each of the
TSAR libraries, described in Sections 6.1 and 6.2 above, had a
potential coding complexity of 20^{36} (~10^{47}) and 20^{20}, respectively,
and since ~10^{14} molecules were used in each transformation
15   experiment, each member of these TSAR libraries should be unique.
After plate amplification the library solution or stock has 10^4
copies of each member/mL.

It was observed that very few (<10%) of the inserted
oligonucleotide sequences characterized so far in both of the
20   libraries have exhibited deletions or insertions. This is likely
a reflection of the accuracy assembling the oligonucleotides
under the conditions used and the fact that certain types of
mutations (i.e., frame-shifts) would not be tolerated as pIII an
essential protein for phage propagation.

25   In order to determine whether any coding bias existed in
the variant non-predicted peptides expressed by these libraries,
perhaps due to biases imposed *in vitro* during synthesis of the
oligonucleotides or *in vivo* during expression by the reproducing
phage, inserts were sequenced as set forth below.

30   **6.3.1.   Characterization of TSAR-9 Library**

Inserted synthetic oligonucleotide fragments of 23
randomly chosen isolates were examined from the TSAR-9 library.
Individual plaques were used to inoculate I ml of 2XYT broth

- 39 -

containing *E. coli* (DH5αF') cells and the cultures were allowed
to grow overnight at 37°C with aeration. DNA was isolated from
the culture supernatants according to Maniatis, *supra*. Twenty-
three individual isolates were sequenced according to the method
5   of Sanger (Proc. Natl. Acad. Sci. USA (1979) 74:5463-5467) using
as a primer the oligonucleotide 5'-AGCGTAACGATCTCCCG (SEQ ID NO.
99), which is 89 nucleotides downstream of the pIII gene cloning
site of the m663 vector used to express the TSARS.

        Nucleotide sequences and their encoded amino acid sequences
10   were analyzed with the MacVector computer program (IBI, New
Haven, CT). The Microsoft EXCEL program was used to evaluate
amino acid frequencies. Such analyses showed that the nucleotide
codons coding for and hence most amino acids, occurred at the
expected frequency in the TSAR-9 library of expressed proteins.
15   The notable exceptions were glutamine and tryptophan, which were over-
and under-represented, respectively.

        It is of interest to note the paucity of TAG stop codons in
the inserts, i.e., only 2 of ~200 isolates characterized
contained a TAG stop codon. About half [1-(47/48)^{36}] of the
20   phage inserts were expected to have at least one TAG codon in
view of the assembly scheme used. However, most of the TAG-
bearing phage appear to have been lost from the library, even
though the bacterial host was *supE*. This may be a consequence of
suppression being less than 100% effective.
25   The amino acids encoded by the inserted double stranded
synthesized oligonucleotide sequences, excluding the fixed PG-
encoding centers, were concatenated into a single sequence and
the usage frequency determined for each amino acid using the
Microsoft EXCEL program. These frequencies were compared to that
30   expected from the assembly scheme of the oligonucleotides, and
the divergence from expected values represented by the size of
the bars above and below the baseline. Chi square analysis was
used to determine the significance of the deviations. The
majority of amino acids were found to occur at the expected
35   frequency, with the notable exceptions that glutamine and

- 40 -

tryptophan were somewhat over- and under-represented, respectively. Thus, except for the invariant Pro-Gly, any position could have any amino acid; hence, the sequences are unpredicted or random.

5   6.3.2.   Characterization of TSAR-12 Library
                    Approximately 10 randomly chosen inserted
oligonucleotides from the TSAR-12 library were examined by DNA
sequencing as described above in Section 6.3.1. The isolates
were chosen at random from the TSAR-12 library and prepared for
10   sequencing, as were the TSAR-9 isolates. Analysis showed that
except for the invariant Gly any position could have any amino
acid; hence, the sequences are unpredicted or random.

          6.4. Preparation of R8C Library
                    Referring now to Figure 3, two oligonucleotides were
15   synthesized on an Applied Biosystems Model 380a machine with the
      sequence 5'-
      TGACGTCTCGAGTTGTNNKNNKNNKNNKNNKNNKNNKNNKTGTGGATCTAGAAGGATC-3'
      (SEQ ID NO:31) and 5'-GATCCTTCTAGATCC-3' (SEQ ID NO:32), where N
      is an equimolar ratio of deoxynucleotides A, C, G, and T, and K
20   is an equimolar ratio of G and T. Fifty pmol of each
      oligonucleotide was incubated at 42 °C for 5 min, then 37 °C for
      15 min, in 50 µL of Sequenase™ buffer (U.S. Biochemicals,
      Cleveland, OH) with 0.1 µg/µL acetylated BSA, and 10 mM DTT.
      After annealing, 10 units of Sequenase™ (U.S. Biochemicals) and
25   0.2 mM of each dNTP were added and incubated at 37 °C for 15 min.
      The sample was then heated at 65 °C for 2 hr, digested with 100
      units of both Xho I and Xba I (New England BioLabs, Beverly, MA),
      phenol extracted, ethanol precipitated, and resolved on a 15%
      non-denaturing polyacrylamide gel. The assembled, digested
30   fragment was gel purified prior to ligation. The vector, m663
      (Fowlkes, D. et al. Biotech. (1992) 13:422-427), was prepared by
      digestion with Xho I and Xba I, calf alkaline phosphatase
      (Boehringer Mannheim, Indianapolis, IN) treatment, phenol

extracted, and purified by agarose gel electrophoresis. To
ligate, 20 μg vector was ·combined with 0.2 μg insert in 3 mL with
T4 DNA ligase (Boehringer Mannheim), according to the
manufacturer. After removal of the protein and buffer by phenol

5   extraction and ethanol precipitation, the ligated DNA was
electroporated into XL1-Blue *E. coli* (Stratagene, San Diego, CA)
and plated for eight hours at 37 °C. To recover the recombinant
phage, the top agar was collected with a spatula, mixed with an
equal volume of 100 mM NaCl, 10 mM $MgCl_2$, and 50 mM Tris-HCI

10   (pH7.5), and disrupted by two passes through an 18-gauge syringe
needle. The bacterial cells were removed by centrifugation, and
phage particles were collected by polyethylene glycol
precipitation and stored at -70 °C in 25% glycerol. The library
had 10^8 total recombinants and a working titer of 6 x 10^{13} pfu/mL.

15   Members of the library were checked for inserts by the
polymerase chain reaction (Saiki, et al. <u>Science</u> (1988) 239:487-
491). Individual plaques on a petri plate were touched with a
sterile toothpick and the tip was stirred into 2xYT with F⁺ *E.
coli* bacteria and incubated overnight at 37 °C with aeration.

20   Five microliters of the phage supernatant were then transferred
to new tubes containing buffer (67 mM Tris-HCl, pH 8.8/10 mM β-
mercaptoethanol/16.6 mM ammonium sulfate/6.7 mM EDTA/50 μg bovine
serum albumin per mL), 0.1 mM deoxynucleotide triphosphates, and
1.25 units of *Taq* DNA polymerase (Boehringer Mannheim,

25   Indianapolis, IN) with 100 pmoles of oligonucleotide primers.
The primers flanked the cloning site in gene III of m663 (5'-
TTCACCTCGAAAGCAAGCTG-3' (SEQ ID NO:100) and
5'-CCTCATAGTTAGCGTAACG-3' (SEQ ID NO:101)). The assembly
reactions were incubated at 94 °C for 1 min, 56 °C for 2 min, and

30   72 °C for 3 min; this cycle was repeated 24 times. The reaction
products were then resolved by electrophoresis on a NuSieve 2.0%
agarose gel (FMC, Rockland, ME). Gels revealed that for 20
plaques tested, all were recombinant and had single inserts of
the expected size.

Based on the sample size of the library, it was anticipated that 100% of the recombinants had single inserts. However, all of the SH3-binding phage isolated from the R8C library had double-inserts. Such phage are presumed rare (i.e., <5%) within
5 the library, yet because the SH3-binding peptide appears to need to be linear they were selected for by our screening methods. Most likely they were formed during the generation of the library; one scenario is that the inserts ligated together to form head-to-head dimers and that they were subsequently cloned
10 into m663 DNA by ligation with the vector's *Xho I* sticky end and by illegitimate ligation with the vector's *Xba I* site (see, Figure 4).

6.5. Preparation Of Target-Coated Microtiter Wells
6.5.1.   Preparation Of GST-SH3 Fusion Proteins
15   The preparation of Src-GST fusion protein was first described by Smith and Johnson, in <u>Gene</u> (1988) 67:31, the disclosure of which is incorporated by reference herein. Briefly, pGEX-derived (Pharmacia, Piscataway, NJ) constructs expressing GST fusion proteins containing the SH3 domains of Src,
20 Grb2, Crk, Abl, or PLCγ were obtained from Dr. Channing Der (University of North Carolina at Chapel Hill); a construct expressing the SH3 domain of Yes was obtained from Dr. Marius Sudol (Rockefeller University). The use of the pGEX bacterial expression vector for the production of GST-SH3 fusion proteins
25 is well-known to those in the art. See, e.g., Cicchetti, P. et al., in <u>Science</u> (1992) 257:803-806. Briefly, the coding region for a particular SH3 domain can be fused in-frame at the Bam HI site of pGEX-2T. Thus, fusion proteins were prepared as per the manufacturer's instructions, and quantified by Coomassie Blue
30 staining of SDS-polyacrylamide gels. Microtiter wells were coated with 5-20 μg GST-SH3 fusion protein in 100 mM $NaHCO_3$, pH 8.5, blocked with 100 mM $NaHCO_3$ (pH 8.5) 1% BSA, and washed. All washes consisted of five applications of 1XPBS, 0.1% Tween 20, 0.1% BSA (Buffer A). Where appropriate, the amount of protein

- 43 -

bound to each well was quantified with an anti-GST antibody-based
ELISA (Pharmacia, Piscataway, NJ), and with a GST-binding phage,
isolated during the course of this work.

6.5.2.   Coating of Microtiter Wells

5   Bacterially expressed Src SH3 glutathione-S-
transferase (Src-GST) fusion protein was purified from bacterial
lysates using glutathione agarose 4B (Pharmacia), according to
the manufacturer's instructions. Bound Src-GST fusion protein
was eluted from the glutathione agarose with 10 mM glutathione in
10   PBS. Microtiter wells were then coated with Src-GST fusion
protein (1-10 μg/well, in 50 mM NaHCO$_3$, pH 8.5) overnight at 4
°C. To block non-specific binding of phage, 100 μL 1% BSA in 100
mM NaHCO$_3$, pH 8.5, was added to each well and allowed to incubate
at room temperature for 1 hour. The wells were then washed five
15   times with 200 μL PBS, 0.1% Tween 20, 0.1% BSA (Buffer A).

6.6. Biopanning And Subsequent Characterization Of Phage-Displayed Random Peptide Libraries With Src-GST Fusion Protein As Target Molecule

6.6.1.   Isolation of Src SH3-Binding Phage

20   Library screens were performed as previously
described. Kay, B.K., et al., in Gene (1993) 128:59-65.
Briefly, 1 X 10^{11} pfu TSAR 9, TSAR 12, or R8C phage in Buffer A
were incubated in a Src SH3-GST-coated well for 2 hours. The
wells were washed, and bound phage were eluted with 100 μL 50 mM
25   glycine·HCl (pH 2.2), transferred to a new well, and neutralized
with 100 mL 200 mM NaHPO$_4$ (pH 7.0). Recovered phage were used to
infect 1 x 10^9 DH5αF' E. coli cells in 20 mL 2xYT; the infected
cells were grown overnight, resulting in a 1000- to 10,000-fold
amplification of phage titer. Amplified phage were panned twice
30   more, as above, excepting the amplification step. Binding phage
recovered after the third round of panning were plated at a low
density on a lawn of DH5αF' E. coli cells to yield isolated
plaques for clonal analysis. Isolated plaques were used to
produce small cultures from which phage stocks and DNA were

recovered for phage binding experiments and dideoxy sequencing
(Sanger, F., et al., in Proc. Natl. Acad. Sci. USA (1977)
74:5463-5467), respectively. Clones were confirmed as binding
the SH3 domain by applying equal titers of phage to wells

5   containing Src SH3-GST or GST alone, and titering the number of
eluted particles from each well, or detecting bound phage with an
anti-phage antibody-based ELISA (Pharmacia).

      Indeed, the ability of isolated phage clones to bind to
several SH3 domains derived from a variety of different proteins

10   can be investigated by the manner described above. GST-SH3
fusion proteins containing SH3 domains from a variety of
different proteins are bound to microliter wells. An aliquot of
the aforementioned phage stocks (50 μL) is introduced into wells
containing the different GST-SH3 fusion proteins. After room

15   temperature incubation for 1-2 hours, the liquid contents of the
microtiter plates are removed, and the wells are washed 5 times
with 200 μL Buffer A. Bound phage are eluted with 100 μL 50 mM
glycine (pH 2.2), transferred to a new well, and neutralized with
100 μL 200 mM NaHPO$_4$ (pH 7.0). The phage are diluted 10^{-3}- to 10^{-6}-

20   ^{6}-fold, and aliquots are plated onto lawns of DH5αF' *E. coli*
cells to establish the number of plaque forming units in the
output sample. From these experiments, the relative specificity
of different Src SH3 binding clones for SH3 domains derived from
other proteins is determined.

25   6.6.2.   **Phage ELISA and Nucleotide Sequencing**
        To evaluate the binding of isolates to
various targets proteins, enzyme-linked-immuno-assays (ELISA)
were also performed. Bacterial cultures were infected with phage
isolates and cultured overnight in 2XYT at 37 °C. The cells were

30   spun down and 25 mL of supernatant was added to microtiter plate
wells coated with 50 mL of protein (1 mg/mL in 100 mM NaHCO$_3$, pH
8.4; overnight at 4 °C or for a few hours at room temperature)
and blocked (1 mg/mL BSA in 100 mM NaHCO$_3$, pH 8.4; for about one
hour). The phage are incubated in the well with 25 mL of

- 45 -

PBS-0.1% Tween 20 at RT for 2 hr. The wells are then washed
multiple times over 30 minutes. To each well is added 50 μL of
polyclonal anti-phage antibody conjugated to horseradish
peroxidase. The antibody is diluted 1:3000 in BSA-PBS-Tween 20;
5   it was obtained from Pharmacia (Piscataway, NJ; catalog number
27-9402-01). After 30 minutes, the wells are washed again with
BSA-PBS-Tween 20 for ~20 minutes. Finally, 100 μL of ABTS reagent
(Pharmacia, with H_2O_2) are added to each well for the development
of color. Plates are read with a plate reader (Molecular Devices,
10   Menlo Park, CA) at 405 nm wavelength.

The nucleotide sequence of the relevant segments of the Src
SH3 binding clones (or phage clones that bind to SH3 domains of
other proteins) were sequenced using standard methods. Sanger,
F., et al., in Proc. Natl. Acad. Sci. USA (1977) 74:5463-5467.
15   The oligo primer 5'-AGCGTAACGATCTAAA-3' (SEQ ID NO:102) was used,
which is 89 nucleotides downstream of the gene III cloning site
of M13 m666. The nucleotide sequences were analyzed with the
MacVector computer program (IBI, New Haven, CT, USA). From this
nucleotide sequence information the primary sequence of each Src
20   SH3 binding peptide was deduced. The corresponding synthetic
peptides were then prepared by techniques well known in the art
with or without flanking sequences. Indeed, these synthetic
peptides have been shown to bind to SH3 domain targets, with
those possessing the phage flanking amino acid residues
25   exhibiting greater binding affinity.

6.7 *In Vitro* Peptide Binding Assays

Peptides were obtained from Research Genetics
(Birmingham, AL), Chiron Mimotopes (Victoria, Australia), or
synthesized by conventional techniques by Dr. J. Mark Carter of
30   Cytogen Corporation (Princeton, NJ). Peptide purity was assessed
by HPLC and/or mass spectrometry. Biotinylated peptides were
synthesized with either a KSGSG (SEQ ID NO:103) or a GSGS (SEQ ID
NO:104) peptide linker (a spacer) between the biotin and the N-
terminus of the peptide. Binding experiments were performed as

- 46 -

above, excepting the use of 10 μM peptide instead of phage.
Bound biotinylated peptide was detected with streptavidin
conjugated to alkaline phosphatase (Sigma Chemical Co., St.
Louis, MO).   After one hour incubation period at room
5   temperature, the wells were washed, and a solution of 3 mM p-
nitrophenyl-phosphate (US Biochemicals, Cleveland, OH) in 50 mM
NaCO$_3$ (pH 9.8), and 50 mM MgCl$_2$ was added and color allowed to
develop.   Signals were read with an ELISA plate reader (Molecular
Devices, Menlo Park, CA) at 405 nm wavelength.   Binding
10   experiments were performed in triplicate.   The results are
presented in Figures 7 and 8.

6.8. Peptide Competition of GST-SH3 Affinity Precipitations of Cell Lysates

Labeled proteins are prepared by incubating a culture
15   of HeLa cells overnight with $\geq$100 μCi/mL ^{35}S-methionine.   The
cells are then washed and lysed with mild detergent.   This
mixture of radioactive proteins is incubated with Src-GST fusion
protein that has been immobilized on glutathione-linked Sepharose
beads (Pharmacia, Piscataway, NJ).   After several hours of
20   tumbling, the beads are pelleted gently by low-speed
centrifugation, and the supernatant is discarded.   The beads are
then resuspended into a slurry in PBS-0.1% Tween 20, pelleted,
and washed several additional times.   Finally, a 2% SDS solution
is added to the sample, which is then boiled at 100 °C for 3
25   minutes.   Afterward, the sample is centrifuged, and the
supernatant loaded on a 10% polyacrylamide SDS gel for
electrophoresis.   After the proteins have been resolved, the gel
is fixed, dried down, and exposed to X-ray film for
autoradiography or phosphor plates for scanning by a Molecular
30   Dynamics PhosphorImager.

The ability of Src SH3 to bind certain ^{35}S-labeled proteins
is examined for competability with exogenous peptides.   Synthetic
peptides corresponding to phage-displayed inserts and motifs are
added at the time that the lysate is incubated with the Src-GST

- 47 -

fusion protein immobilized on glutathione-linked sepharose beads.
The SH3 binding peptides block binding of all or some of the
labeled proteins while negative control peptides (unrelated
peptide sequences) do not. The amount of competition is

5   quantified and correlated with the amount of added SH3-domain
binding peptides.

       Alternatively, NIH 3T3 cells were grown in Dulbecco's
Modified Eagle Medium (DME) + 10% fetal calf serum (FCS) + 80
μCi/mL Tran^{35}Slabel (ICN), washed with PBS, lysed in RIPA buffer,

10   and pelleted. Supernatant from 1.5 x 10^6 cells was precleared
with 100 μg glutathione-agarose-immobilized GST. The supernatant
was then incubated with 10 μg glutathione-agarose-immobilized
GST-SH3 fusion protein with or without added test peptide in a
final volume of 250 μL. Pelleted beads were washed with 1 mL

15   each of RIPA, RIPA + 1% deoxycholate + 0.1% SDS, and PBS,
resuspended in 50 μL SDS-PAGE sample buffer, boiled, and
subjected to SDS-PAGE (7.5%). Labeled proteins were detected by
phosphorimaging (Molecular Dynamics). The results are presented
in Figure 9.

20   **6.9. Peptide Competition of GST-SH3 Affinity Precipit-
             ations of PI-3' Kinase From Cell Lysates**
             It is possible to follow the precipitation of PI-3'
Kinase by Src from cell lysates in the presence or absence of
SH3-binding peptides. HeLa cells are lysed with detergent and

25   the protein mixtures are incubated for several hours with the
Src-GST fusion protein immobilized on glutathione-linked
Sepharose beads. After several hours of tumbling, the beads are
pelleted gently by low-speed centrifugation and the supernatant
is discarded. The beads are then resuspended into a slurry in

30   PBS-0.1% Tween 20, pelleted, and washed several additional times.
Finally, an SDS solution is added to the sample, which is then
boiled at 100 °C for 3 minutes. Subsequently, the sample is
centrifuged, and the supernatant is loaded on a 10%
polyacrylamide SDS gel for electrophoresis. After the proteins

have been resolved, the gel is blotted to nitrocellulose or nylon (i.e., western blot). The filter is then probed with a PI-3' Kinase antibody (monoclonal and polyclonal antibodies are available from Upstate Biotechnology Incorporated, Lake Placid,

5   NY) and an enzyme-linked secondary antibody. The amount of PI-3' Kinase is then quantitated.

The ability of Src SH3 to bind PI-3' Kinase is examined for competability with exogenous peptides. Synthetic peptides corresponding to phage-displayed inserts and motifs are added at

10 the time that the lysate is incubated with the Src-GST fusion protein that has been immobilized on glutathione-linked sepharose beads. Ten-fold and one hundred-fold molar excess of peptides are used relative to SH3 proteins. The SH3 binding peptides block binding of the PI-3' Kinase as detected on western blots

15 while negative control peptides (unrelated peptide sequences) do not. The amount of competition is quantified and correlated with the amount of added SH3-domain binding peptides.

6.10.   In Vivo Association Of SH3-Binding Peptides With SH3-Domains Of Proteins

20   To demonstrate association of the SH3-binding peptides with SH3-domains of proteins inside cells, the SH3-binding peptides are tagged and localized in cells. For example, Bar-Sagi et al., in Cell (1993) 74:83-91, have shown that SH3-binding proteins localize to the cytoskeleton when expressed in

25 cells. Thus, the SH3 domain-binding peptides of the present invention can serve as cellular targetting signals (e.g., to the cytoskeleton). Accordingly, the peptides are tagged with biotin and, subsequently, injected into cells. Alternatively, one can transfect into cells a recombinant plasmid that expresses a

30 fusion protein comprising of the SH3-binding peptide and the green fluorescent protein (GFP, Chalfie et al., in Science (1994) 263:802-805). The location of the biotinylated peptide or the GFP fusion protein is then assayed with FITC-labeled streptavidin in paraformaldehyde-fixed cells or by direct fluorescence in

- 49 -

living cells, respectively. Localization of the SH3-binding
peptides to the cytoskeleton demonstrates that the SH3-binding
peptides can bind SH3-domain proteins *in vivo*. In addition,
focal adhesions, which are rich in Src, are also sites of

5   potential subcellular localization of SH3-binding peptides.

Thus, NIH 3T3 fibroblasts were cultured *in vitro* on glass
coverslips coated with fibronectin. After two days of growth at
37 °C, the cells were fixed for one hour at room temperature in
the presence of 2% paraformaldehyde (pH 7.5). The coverslips

10   were washed with PBS-0.1% Tween 20 several times to remove the
fixative. Next, the coverslips were dipped into acetone (chilled
at -20 °C) for approximately 20 seconds and allowed to air-dry.
The coverslips were washed again with PBS-0.1% Tween 20,
containing BSA (1 mg/mL), and incubated for 2 hours at room

15   temperature with different biotinylated peptides in PBS-0.1%
Tween 20. The coverslips were washed and then incubated with 1
mg/mL streptavidin-Cy3 (Jackson Immunoresearch Co., West Grove,
PA) for 1 hour at room temperature. Finally, the coverslips were
washed in PBS-0.1% Tween 20, mounted in a glycerol solution on a

20   glass slide, and viewed with a Nikon Optiphot epifluorescence
microscope and a 60x oil immersion lens.

The results are presented in Figure 11, in which panel A
displays cells stained with the conjugate biotin-spacer-
VLKRPLPIPPVTR (SEQ ID NO:64); panel B exhibits cells stained with

25   the conjugate, biotin-spacer-GILAPPVPPRNTR (SEQ ID NO:63); panel
C shows cells stained with the long consensus peptide, biotin-
spacer-RSTPRPLPPLPTTR (SEQ ID NO:67); and panel D shows cells
stained with the proline-rich vinculin peptide conjugate, biotin-
spacer-LAPPKPPLPEGEV (SEQ ID NO:70). The "spacer" sequence is

30   KSGSG (SEQ ID NO:103). As shown in Figure 11, the panels in
which SH3 domain-binding peptides were used present a bright
display of fluorescence activity that is in sharp contrast to the
relatively "dark" features of panel D (non-SH3 domain binding
vinculin segment). These results demonstrate further the ability

35   of the SH3 domain-binding peptides of the present invention to

- 50 -

localize to protein targets (e.g., Src and Src-related proteins) within cells and provide an image thereof.

6.11.   In Vivo Modulation Of Src In Oocytes With SH3-Binding Peptides

When *Xenopus laevis* oocytes are injected with mRNA encoding deregulated Src, there are *dramatic* cytological and biochemical changes in the oocyte (Unger, T.F. and Steele, R.E., in Mol. Cell. Biol. (1992) 12:5485-5498). The applicants have obtained the plasmid for generating Src mRNA, which is available from Dr. Robert Steele (University of California at Irvine). Synthetic SH3-binding peptides are injected into oocytes that have been previously injected with Src mRNA. The state of the cytoskeleton is inspected visually by observing the arrangement of cortical pigment granules under a dissecting microscope. The state of phosphorylation of several proteins is examined by western blotting with an anti-phosphotryosine monoclonal antibody (4G10; Upstate Biotechnology Incorporated), as described in Unger and Steele, above.

6.12.   Progesterone-induced *X. laevis* Oocyte Maturation

Segments of adult ovary were removed surgically and incubated in 0.1% collagenase type D (Boehringer Mannheim, Indianapolis, IN) in Ca^{2+}-free OR2 (82.5 mM NaCl, 2.5 mM KCl, 1.0 mM $MgCl_2$, 1.0 mM Na_2HPO_4, 5.0 mM HEPES, and 3.8 mM NaOH, pH 7.6). Oocytes were then washed 3-5 times with OR2 containing 1.0 mM $CaCl_2$ and allowed to recover in OR2 overnight at 18 °C. Stage VI oocytes were injected with 40 nL of 100 mM peptide or water. After injection, the oocytes were placed in OR2 with 2 mg/mL progesterone (Sigma, St Louis, MO) and incubated at 20 °C. Oocytes were scored at hourly time points for germinal vesicle breakdown (GVBD).

Figure 10 presents the results of this experiment. As shown by the graph, oocytes injected with the SH3 domain-binding

- 51 -

peptide VLKRPLPIPPVTR (SEQ ID NO:64) exhibit a faster rate of
progesterone-induced germinal vesicle breakdown relative to
oocytes that had been injected with water or with the proline-
rich vinculin peptide, LAPPKPPLPEGEV (SEQ ID NO:70). These
5   results parallel those obtained by Unger and Steele, *supra*, who
observed a gross alteration in the cortex of oocytes that had
been injected with RNA encoding a deregulated Src variant
("active" Src) versus those injected with RNA encoding the wild-
type Src ("cellular" Src). Also, as shown in Figure 3B of the
10   Unger and Steele article, oocytes injected with deregulated or
active Src RNA matured at a faster rate than oocytes injected
with water or wild-type Src RNA.

        The present results obtained with Src SH3 domain-binding
peptides suggest that these peptides modulate the biochemical
15   activity of "cellular" Src; in particular, it is proposed that at
least some of the Src SH3 domain-binding peptides of the present
invention upregulate the biochemical activity of "cellular" Src,
which may be downregulated or inhibited in its normal state.
Hence, the administration of the SH3 domain-binding peptides of
20   the present invention can constitute a novel method of modulating
the activity of Src or Src-related proteins. Specifically,
certain of these peptides are able to activate Src-family
proteins.

       6.13.   In Vivo Antagonism Of Src In Src Transformed
25   **Cells With SH3-Binding Peptides**
                  The coding regions for SH3-binding peptides are
cloned into vectors that direct their expression in animal cells.
A bipartite gene is constructed, encoding a protein with *c-myc*
epitope and SH3-binding peptide, which is transcribed from a
30   strong constitutive promoter (e.g., SV40, CMV, HSV TK,
calmodulin). The vector is introduced into either normal or Src-
transformed cells via transfection (e.g., electroporation,
calcium phosphate, liposomes, DEAE dextran). Transfected cells
express the bipartite gene transiently in culture. To create

                                 - 52 -

stable transformed cell lines, the vector carries a selectable
marker (e.g., neomycin resistance) or transfection is performed
in the presence of excess plasmid carrying a selectable marker
(e.g., neomycin resistance) and cells selected for the marker.

5   Transfected cells are stained by immunofluorescence to detect
expression of the bipartite protein. The hybridoma 9E10 secretes
a monoclonal antibody that is highly specific for the *c-myc*
epitope (EQKLISEEDLN [SEQ ID NO:105]; see, Evan, G.A. et al., in
Mol. Cell. Biol. (1985) 5:3610-3616). This antibody is used in

10   immunofluorescence experiments to demonstrate that the bipartite
protein is expressed inside the cells, and in some cases,
localized to subcellular structures enriched in SH3 domain
bearing proteins.

There are several controls used in these experiments.

15   First, cells are transfected with vectors that do not have the
SH3-binding peptide coding region. Second, normal (non-
transformed) cells are transfected. Third, cells transformed by
oncogenes other than Src are used in the transfection
experiments. Fourth, cells are stained with other monoclonal

20   antibodies that do not recognize the *c-myc* epitope.

Transfected cells are examined for any changes in cell
shape, behavior, and metabolism as a consequence of expressing
the SH3 binding peptides. Cell shape is examined by phase
contrast microscope at several times after transfection; in

25   particular, the flatness of the cells, their adhesion to the
substrate, and the degree of cell ruffling are monitored. Cell
division rates, cell migration, and contact inhibition are also
observed over time. Finally, the amount of phosphorylated
tyrosine in transfected cells is quantitated by phosphoaminoacid

30   analysis and with an anti-phosphotryosine monoclonal antibody
(4G10; Upstate Biotechnology Incorporated) in western blotting
experiments.

6.14 Preparation of PXXP Biased Peptide Libraries

The preparation and characterization of preferred phage-displayed random peptide libraries have been described above in Sections 6.1 - 6.4.

5   Using procedures similar to those described in these sections, oligonucleotide inserts were contructed according to the schematic provided in FIG. 12. The inserts were then cloned into the mBAX vector, and the biased peptide libraries were expressed as described previously.

10   The mBAX vector was created in the Kay Laboratory by generating cloning sites in gene III of the M13mp18 vector (Messing, J. (1991). Cloning in M13 phage or how to use biology at its best. *Gene* 100, 3-12) in the manner of Fowlkes et al. (1992)(*Biotechniques* 13, 422-427). The mBAX vector displays a

15 peptide sequence at the N-terminus of the mature pIII protein that encodes the epitope for the mouse monoclonal antibody 7E11 (see, FIG. 13); it includes the stop codon TAG in the coding region, which is suppressed in *E. coli* carrying suppressor tRNA gene mutations known as *supE* or *supF*. There are no other stop

20 codons in the mBAX genome. The mBAX vector also carries a segment of the alpha fragment of ß-galactosidase; bacterial cells expressing the omega fragment of ß-galactosidase convert the clear XGal substrate into an insoluble blue precipitate. The plaques appear blue.

25   Recombinant mBAX molecules can be distinguished easily from non-recombinant molecules due to the TAG codon in the *XhoI* - *XbaI* segment in gene III of mBAX. When recombinants are generated by replacing the *Xho* I - *Xba* I fragment with oligonucleotides encoding random peptides, the recombinants can be grown in

30 bacteria with (i.e., DH5αF') or without (i.e., JS5) suppressor tRNA mutant genes. On the other hand, the non-recombinant mBAX molecules fail to produce plaques on bacterial lawns where the bacteria (i.e., JS5) lack such suppressor genes. This is because in JS5, the TAG codon serves as a stop codon to yield a truncated

35 pIII molecule during translation; since pIII is an essential

protein component of viable M13 viral particles, no plaques will
form.

6.14.1   Discussion

The use of second generation or biased peptide libraries,
5   which fix all or part of the RPLPPLP consensus motif for SH3
binding peptides and randomize flanking residues, has defined
additional sequence residues exhibiting selective SH3 binding.

Tables 1-5, below, list the relevant amino acid sequences
obtained from the biased peptide library for each set of SH3
10   domain binding peptides. The underscored amino acid residues
indicate the fixed positions. Also, indicated for each set of
new binders is a "consensus" sequence, which seeks to include the
additional features gleaned from the new binding peptides. The
symbol "ϕ" represents a hydrophobic residue.

TABLE 1   CORTACTIN SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.CORT.M1/2/3.PP	SSLLGPPVPPKPQTLFSFSR	107
PXXP.CORT.M4.PP	SRLGEFSKPPIPQKPTWMSR	108
PXXP.CORT.N2.PP	SRTERPPLPQRPDWLSYSSR	109
PXXP.CORT.N3.PP.INC	SREPDWLCPNCPLLLRSDSR	110
PXXP.CORT.01/2/3.PP	SSSSHNSRPPLPEKPSWLSR	111
PXXP.CORT.04.PP	SRLTPQSKPPLPPKPSAVSR	112
CONSENSUS	KPPϕPxKPxW R	113

- 55 -

TABLE 2   NCK SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.NCK.Q1/4.PP	SSLGVGWKPLPPMRTASLSR	114
PXXP.NCK.Q2/3.PP.INC	SSVGFADRPRPPLRVESLSR	115
PXXP.NCK.R1.PP.INC	SSAGILRPPEKPXRSFSLSR	116
PXXP.NCK.R2.PP	SSPYTGDVPIPPLRGASLSR	117
PXXP.NCK.R3.PP	SSLMGSWPPVPPLRSDSLSR	118
PXXP.NCK.R4.PP	SSIGEDTPPSPPTRRASLSR	119
PXXP.NCK.S1/4.PP	SRSLSEVSPKPPIRSVSLSR	120
PXXP.NCK.S2.PP.INC	SSVSEGYSPPLPPRSTSLSR	121
PXXP.NCK.S3.PP	SSSFTLAAPTPPTRSLSLSR	122
PXXP.NCK.T1.PP	SSPPYELPPRPPNRTVSLSR	123
PXXP.NCK.T2.PP	SRVVDGLAPPPPVRLSSLSR	124
PXXP.NCK.T3.PP.INC	SSLGYSGAPVPPHRxSSLSR	125
PXXP.NCK.T4.PP	SSISDYSRPPPPVRTLSLSR	126
CONSENSUS	ϕxxxxxxPxPPϕRSxSL T	127

TABLE 3   ABL SH3 BINDING PEPTIDES

		SEQ. ID NO.
PXXP.ABL.G1/2.PP	SRGPRWSPPPVPLPTSLDSR	128
PXXP.ABL.G3/4.PP	SSPPDYAAPAIPSSLWVDSR	129
PXXP.ABL.H1/3/4.PP	SSPPHWAPPAPPAMSPPISR	130
PXXP.ABL.H2.PP.INC	SSDRCWECPPWPAGGQRGSR	131
PXXP.ABL.I1/2/3.PP	SSPPKFSPPPPPYWQLHASR	132
PXXP.ABL.I4.PP	SSPPSFAPPAAPPRHSFGSR	133
PXXP.ABL.J1.PP	SSAPKKPAPPVPMMAHVMSR	134
PXXP.ABL.J2.PP.INC	SSPTYPPPPPPDTAKGASR	135
PXXP.ABL.J3.PP.INC	SSPPXXXPPPIPNSPQVLSR	136
PXXP.ABL.J4.PP	SSPPTWTPPKPPGWGVVFSR	137
PXXP.ABL.L1.PP	SSAPTWSPPALPNVAKYKSR	138
PXXP.ABL.L2/3.PP	SSIKGPRFPVPPVPLNGVSR	139
PXXP.ABL.L4.PP	SSPPAWSPPHRPVAFGSTSR	140
CONSENSUS	PPxWxPPPφP	141

TABLE 4   PLCG SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.PLCG.P1.PP	SSMKVHNFPLPPLPSYETSR	142
PXXP.PLCG.P2.PP	SRVPPLVAPRPPSTLNSLSR	143
PXXP.PLCG.PE.PP.INC	SSLYWQHGPDPPVGAPQLSR	144
PXXP.PLCG.P4.PP	SSHPLNSWPGGPFRHNLSSR	145

- 57 -

TABLE 5   SRC SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.SRC.A1.PP	SSRALRVRPLPPVPGTSLSR	146
PXXP.SRC.A2.PP	SSFRALPLPPTPDNPFAGSR	147
PXXP.SRC.A3.PP	SRDAPGSLPFRPLPPVPTSR	148
PXXP.SRC.A4.PP	SSISQRALPPLPLMSDPASR	149
PXXP.SRC.B1.PP	SSPAYRPLPRLPDLSVIYSR	150
PXXP.SRC.B2/3/PP	SSFINRRLPALPPDNSLLSR	151
PXXP.SRC.B4.PP	SRLTGRPLPALPPPFSDFSR	152
PXXP.SRC.C1.PP	SRMKDRVLPPIPTVESAVSR	153
PXXP.SRC.C2.PP.INC	SSLYSAIAPDPPPRNSSSSR	154
PXXP.SRC.C3.PP	SSLASRPLPLLPNSAPGQSR	155
PXXP.SRC.D1.PP	SSLTSRPLPDIPVRPSKSSR	156
PXXP.SRC.D2.PP.INC	SSLKWRALPPLPETDTPYSR	157
PXXP.SRC.D3.PP	SSNTNRLPPPTPDGLDVRSR	158
PXXP.SRC.D4.PP	SSLQSRPLPLPPQSSYPISR	159
CONSENSUS	RPLPPLP	160

It should be apparent to one of ordinary skill
that many other embodiments of the present invention
can be contemplated beyond the preferred embodiments
described above but which other embodiments
5   nevertheless fall within the scope and spirit of the
present invention. Hence, the present invention should
not be construed to be limited to the preferred
embodiments described herein, which serve only to
illustrate the present invention, but only by the
10   claims that follow.
       Also, numerous references are cited throughout the
specification. The complete disclosures of these
references are incorporated by reference herein.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: KAY, BRIAN K.
             SPARKS, ANDREW B.
             THORN, JUDITH M.
             QUILLIAM, LAWRENCE A.
             DER, CHANNING J.

   (ii) TITLE OF INVENTION: Src SH3 BINDING PEPTIDES AND METHODS OF
        ISOLATING AND USING SAME

  (iii) NUMBER OF SEQUENCES: 106

   (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Pennie & Edmonds
        (B) STREET: 1155 Avenue of the Americas
        (C) CITY: New York
        (D) STATE: New York
        (E) COUNTRY: U.S.A.
        (F) ZIP: 10036-2711

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

   (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

 (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: S. Leslie Misrock
        (B) REGISTRATION NUMBER: 18,872
        (C) REFERENCE/DOCKET NUMBER: 1101-192

   (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (212) 790-9090
        (B) TELEFAX: (212) 869-9741/8864
        (C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8
        (D) OTHER INFORMATION: /note= "hydrophobic residue (such
            as Pro or Leu)"

- 59 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Xaa Pro Xaa Xaa Pro Pro Pro Xaa Xaa Pro
1   5   10

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1..3
        (D) OTHER INFORMATION: /note= "Xaa = any residue other
            than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "X = any residue other
            than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8..9
        (D) OTHER INFORMATION: /note= "X = any residue other
            than Cys"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Xaa Xaa Xaa Pro Pro Xaa Pro Xaa Xaa
1   5

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "X tends to be hydrophobic"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8..9
        (D) OTHER INFORMATION: /note= "X tends to be hydrophobic"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Xaa Leu Pro Pro Arg Pro Xaa Xaa
1   5

(2) INFORMATION FOR SEQ ID NO:4:

     (i) SEQUENCE CHARACTERISTICS: ·
          (A) LENGTH: 9 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

     Arg Lys Leu Pro Pro Arg Pro Arg Arg
     1   5

(2) INFORMATION FOR SEQ ID NO:5:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 10 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

     Val Pro Pro Pro Val Pro Pro Arg Arg Arg
     1   5   10

(2) INFORMATION FOR SEQ ID NO:6:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 7 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 2
          (D) OTHER INFORMATION: /note= "tends to be Pro"

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 3
          (D) OTHER INFORMATION: /note= "hydrophobic residue"

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 5
          (D) OTHER INFORMATION: /note= "tends to be Pro"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

     Xaa Xaa Xaa Pro Xaa Xaa Pro
     1   5

(2) INFORMATION FOR SEQ ID NO:7:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 9 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: peptide

  (ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 2
     (D) OTHER INFORMATION: /note= "X = any residue other
         than Cys"

  (ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 8..9
     (D) OTHER INFORMATION: /note= "X = any residue other
         than Cys"

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Arg Xaa Leu Pro Pro Arg Pro Xaa Xaa
1   5

(2) INFORMATION FOR SEQ ID NO:8:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 9 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: peptide

  (ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 9
     (D) OTHER INFORMATION: /note= "hydrophobic residue"

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Arg Xaa Leu Pro Pro Leu Pro Arg Phe
1   5

(2) INFORMATION FOR SEQ ID NO:9:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 7 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Arg Pro Leu Pro Pro Leu Pro
1   5

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS: ·
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "any residue other than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5..6
        (D) OTHER INFORMATION: /note= "any hydrophobic residue"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 9
        (D) OTHER INFORMATION: /note= "any hydrophilic residue
            except Cys"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Arg Xaa Leu Pro Xaa Xaa Pro Xaa Xaa
1   5

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Arg Pro Leu Pro Pro Leu Pro Thr Ser
1   5

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (v) FRAGMENT TYPE: C-terminal

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Xaa Xaa Xaa Xaa
1

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (v) FRAGMENT TYPE: C-terminal

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Xaa Xaa Xaa Xaa Xaa
1   5

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (v) FRAGMENT TYPE: N-terminal

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Xaa Xaa Xaa Xaa
1

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1..3
        (D) OTHER INFORMATION: /note= "any residue other than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "any residue other than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8..9
        (D) OTHER INFORMATION: /note= "any hydrophobic residue"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 11..13
        (D) OTHER INFORMATION: /note= "any residue other than Cys"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

        Xaa Xaa Xaa Arg Xaa Leu Pro Xaa Xaa Pro Xaa Xaa Xaa
        1   5   10

(2) INFORMATION FOR SEQ ID NO:16:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 42 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

        Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        1   5   10   15

        Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                    20   25   30

        Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg
                35   40

(2) INFORMATION FOR SEQ ID NO:17:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 69 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGCTCGAGNN NBNNBNNBNN BNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNNBNNBNNB   60

NNBCCAGGT   69

(2) INFORMATION FOR SEQ ID NO:18:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 68 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGTCCANNVN NVNNVNNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVNN VNNVNNVNNV   60

AGATCTGG   68

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 66 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCGAGNNNBN NBNNBNNBNN BNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNNBNNBNNB   60

CCAGGT   66

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 65 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGTCCANNVN NVNNVNNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVNN VNNVNNVNNV   60

AGATC   65

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 46 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (v) FRAGMENT TYPE: N-terminal

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "Arg or Ser"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1   5   10   15

Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
       20   25   30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Pro Ser Arg Thr
       35   40   45

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS: .
       (A) LENGTH: 49 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: protein

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Ser His Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1               5                   10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Pro Ser Arg
        35                  40                  45

Thr
```

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 27 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: peptide

   (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 2
       (D) OTHER INFORMATION: /note= "Ser or Thr"

   (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 13
       (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

   (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 15
       (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa
1               5                   10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 48 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: both
       (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTTTGTCGAC NNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNGCGGTG   48

(2) INFORMATION FOR SEQ ID NO:25:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 48 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGCCACNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVTG ATCATTTT   48

(2) INFORMATION FOR SEQ ID NO:26:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 43 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CGCCACNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVTG ATC   43

(2) INFORMATION FOR SEQ ID NO:27:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 43 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCGACNNNBN NBNNBNNBNN BNNBNNBNNB NNBNNBNGCG GTG   43

(2) INFORMATION FOR SEQ ID NO:28:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein

        (v) FRAGMENT TYPE: N-terminal

- 68 -

```
(ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 2
     (D) OTHER INFORMATION: /note= "Ser or Thr"

(ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 13
     (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

(ix) FEATURE:
     (A) NAME/KEY: Modified-site
     (B) LOCATION: 15
     (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa
1               5                   10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg Pro Ser Arg Thr
            20                  25                  30
```

```
(2) INFORMATION FOR SEQ ID NO:29:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 34 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: protein

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "Ser or Thr"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 16
        (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 18
        (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"


   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

   Ser His Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
   1               5                   10                  15

   Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg Pro Ser
               20                  25                  30

   Arg Thr
```

(2) INFORMATION FOR SEQ ID NO:30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (v) FRAGMENT TYPE: internal

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1   5

(2) INFORMATION FOR SEQ ID NO:31:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 58 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TGACGTCTCG AGTTGTNNKN NKNNKNNKNN KNNKNNKNNK TGTGGATCTA GAAGGATC   58

(2) INFORMATION FOR SEQ ID NO:32:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 15 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CCTAGATCTT CCTAG   15

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 40 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TCGAGTTGTN NKNNKNNKNN KNNKNNKNNK NNKTGTGGAT   40

(2) INFORMATION FOR SEQ ID NO:34:

       (i) SEQUENCE CHARACTERISTICS: .
            (A) LENGTH: 40 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: DNA (genomic)

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CAACANNMNN MNNMNNMNNM NNMNNMNNMA CACCTAGATC   40

(2) INFORMATION FOR SEQ ID NO:35:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 19 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: protein

       (v) FRAGMENT TYPE: N-terminal

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ser Ser Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Gly Ser Arg Pro
1   5   10   15

Ser Arg Thr

(2) INFORMATION FOR SEQ ID NO:36:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 28 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Ser Ser Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Gly Ser Arg Ser
1   5   10   15

Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Thr Arg
            20   25

(2) INFORMATION FOR SEQ ID NO:37:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 80 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: DNA (genomic)

- 71 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TCGAGTTGTN NKNNKNNKNN KNNKNNKNNK NNKTGTGGAT CTAGATCCAC AVNNVNNVNN   60

VNNVNNVNNV NNVNNACAAC   80

(2) INFORMATION FOR SEQ ID NO:38:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 45 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Ser Ser Phe Asp Gln Gln Asp Trp Asp Tyr Ser Ile Ala Glu Lys Met
1   5   10   15

His Pro Ile Arg Pro Gly Phe Arg Glu Leu Pro Pro Leu Pro Pro Ser
            20   25   30

Arg Ala Ser Phe Gly Gly Gly Ala Ser Arg Pro Ser Arg
        35   40   45

(2) INFORMATION FOR SEQ ID NO:39:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 30 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Ser Thr Asn Val Trp Val Thr Gly Ser Val Ile Ala Arg Gly Ala Gln
1   5   10   15

Ser Arg Pro Leu Pro Ile Pro Pro Glu Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:40:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 30 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ser Thr Ala Pro Trp Gly Leu Arg Val Ala His Glu Gly Gly Val Leu
1   5   10   15

Lys Arg Pro Leu Pro Ile Pro Pro Val Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:41:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 45 amino acids
  (B) TYPE: amino acid
  (D) TOPOLOGY: unknown

 (ii) MOLECULE TYPE: peptide

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
Ser Ser Ser Gly Tyr Val Val Pro Lys Arg Leu Gly Asp Met Arg Glu
1               5               10              15

Tyr Asn Ala His Pro Gly Leu His Val Pro Pro Asn Ser Pro Leu Pro
            20              25              30

Pro Leu Pro Thr His Leu Gln Ser Ser Arg Pro Ser Arg
        35              40              45
```

(2) INFORMATION FOR SEQ ID NO:42:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 44 amino acids
  (B) TYPE: amino acid
  (D) TOPOLOGY: unknown

 (ii) MOLECULE TYPE: peptide

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```
Ser Ser Arg Gly Glu Gly Asn Asn Ile Ile Ser Ser Arg Pro Phe Leu
1               5               10              15

Ser Asn Ser Asp Pro Gly Val Ser Asn Lys Leu Thr Gly Arg Gly Pro
            20              25              30

Leu Pro Pro Leu Pro Asn Asp Ser Arg Pro Ser Arg
        35              40
```

(2) INFORMATION FOR SEQ ID NO:43:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 31 amino acids
  (B) TYPE: amino acid
  (D) TOPOLOGY: unknown

 (ii) MOLECULE TYPE: peptide

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```
Ser Thr Ala Val Ser Phe Arg Phe Met Pro Gly Gly Gly Gly Ala Phe
1               5               10              15

Tyr Ser Thr Arg Pro Val Pro Pro Ile Thr Arg Pro Ser Arg Thr
            20              25              30
```

(2) INFORMATION FOR SEQ ID NO:44:

    (i) SEQUENCE CHARACTERISTICS: .
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ser Thr Ala His Ser Leu Trp Asp Trp Gly Thr Phe Ser Gly Val Ser
1   5   10   15

His Lys Ser Arg Leu Pro Pro Leu Pro Thr Arg Pro Ser Arg Thr
            20   25   30

(2) INFORMATION FOR SEQ ID NO:45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Pro Gly Tyr Ala Arg Ile Val Ser Tyr Arg Phe Arg Ala Leu Pro Ser
1   5   10   15

Pro Pro Ser Ala Ser Arg Pro Ser Arg
            20   25

(2) INFORMATION FOR SEQ ID NO:46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Ser Thr Asn Asp Val Asp Trp Met His Met Trp Asn Ser Gly Gly Pro
1   5   10   15

His Arg Arg Leu Pro Pro Thr Pro Ala Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:47:

    (i) SEQUENCE CHARACTERISTICS: .
        (A) LENGTH: 45 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```
Ser Ser Asp Asn Trp Ala Arg Arg Val His Ala Ser Glu Leu Ile Tyr
1               5                   10                  15

Thr Asp Leu Ser Pro Gly Ile Leu Leu Ala Gln Arg Gln Leu Pro Pro
            20                  25                  30

Thr Pro Gly Arg Asp Pro Ser His Ser Arg Pro Ser Arg
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:48:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 45 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```
Ser Ser Glu Ser Pro Leu Met Tyr Asn Arg Val Gly Ala Leu Gln Ser
1               5                   10                  15

Leu Thr Ser Val Pro Gly Ser Met Met His Phe Ala Leu Gln Arg Arg
            20                  25                  30

Leu Pro Arg Thr Pro Pro Pro Ala Ser Arg Pro Ser Arg
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:49:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 35 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```
Ser Thr Arg Trp Ser His Ser Trp Pro Gly Tyr Val Gly Gly Ala Asn
1               5                   10                  15

Pro Ser Pro Ala Thr Arg Pro Leu Pro Thr Arg Pro Ser Arg Thr Val
            20                  25                  30

Glu Ser Cys
        35
```

(2) INFORMATION FOR SEQ ID NO:50:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 50 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Ser Arg Tyr Asn Asp Leu Gly Thr Arg Pro Val Ser Glu Val Ile Lys
1   5   10   15

Tyr Asp Tyr Phe Pro Gly Tyr Ser Gln His Val Ile Thr Pro Asp Gly
            20   25   30

Ser Tyr Ser Thr Arg Pro Leu Pro Ser Arg Pro Ser Arg Thr Val Glu
        35   40   45

Ser Cys
    50

(2) INFORMATION FOR SEQ ID NO:51:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Pro Gly Arg Leu Leu Leu Pro Ser Glu Pro Arg Thr Phe Tyr Asn Tyr
1   5   10   15

Gly His Asp Ser Arg Pro Ser Arg
            20

(2) INFORMATION FOR SEQ ID NO:52:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Ser Thr Met Tyr Gly Val Ser Trp Leu Ser Ser Gly Ser Gly Gly Ile
1   5   10   15

Leu Ala Pro Pro Val Pro Pro Arg Asn Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:53:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Ser Ser Cys Thr Glu Lys Thr Val Ser Gly Trp Cys Gly Ser Arg Ser
1   5   10   15

Thr Arg Pro Leu Pro Ile Leu Pro Arg Thr Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:54:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Ser Ser Cys Met Leu Pro Thr Asp Gly Trp Gln Cys Gly Ser Arg Ser
1   5   10   15

Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:55:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Ser Ser Cys Asp Gly Thr Gln Phe Arg Leu Asn Cys Gly Ser Arg Ser
1   5   10   15

Thr Asn Arg Pro Leu Pro Met Ile Pro Thr Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:56:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ser Ser Cys Met Gln Gly Gln Ala Gly Leu Lys Cys Gly Ser Arg Ser
1   5   10   15

Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr Arg Pro Ser Arg
            20   25   30

- 77 -

(2) INFORMATION FOR SEQ ID NO:57:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 31 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```
Ser Ser Cys Tyr Arg Glu Lys Asp Thr Trp Gly Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Ser Arg Pro Leu Pro Ser Leu Pro Thr Thr Arg Pro Ser Arg
            20              25                  30
```

(2) INFORMATION FOR SEQ ID NO:58:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 32 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
Ser Ser Cys Leu Phe Glu Gln Gly Ala Gly Thr Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Ser Leu Pro Pro Leu Pro Pro Thr Thr Arg Pro Ser Ser Arg
            20              25                  30
```

(2) INFORMATION FOR SEQ ID NO:59:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 44 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
Ser Ser Cys Asp Thr Gly Arg Ile Ala Pro Gly Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Pro Arg Pro Leu Pro Leu Ile Pro Thr Thr Pro Arg Ser Thr Asn
            20              25                  30

Leu Asn Leu Thr Ser Thr Thr Thr Arg Pro Ser Arg
        35                  40
```

- 78 -

(2) INFORMATION FOR SEQ ID NO:60:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 31 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
Ser Ser Cys Gly Leu Asp Asn Ala Ala Lys Thr Cys Gly Ser Arg Ser
1               5               10              15

Thr Arg Pro Leu Pro Pro Thr Pro Leu Thr Thr Arg Pro Ser Arg
            20              25              30
```

(2) INFORMATION FOR SEQ ID NO:61:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 31 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Ser Ser Cys Ser Arg Ala His Glu Thr Glu Met Cys Gly Ser Arg Ser
1               5               10              15

Thr Arg Pro Gln Pro Pro Pro Pro Ile Thr Thr Arg Pro Ser Arg
            20              25              30
```

(2) INFORMATION FOR SEQ ID NO:62:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 14 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```
Arg Ser Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg
1               5               10
```

(2) INFORMATION FOR SEQ ID NO:63:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Gly Ile Leu Ala Pro Pro Val Pro Pro Arg Asn Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:64:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Val Leu Lys Arg Pro Leu Pro Ile Pro Pro Val Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:65:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Gly Pro His Arg Arg Leu Pro Pro Thr Pro Ala Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:66:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Ala Asn Pro Ser Pro Ala Thr Arg Pro Leu Pro Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:67:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

- 80 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

```
Arg Ser Thr Pro Arg Pro Leu Pro Pro Leu Pro Thr Thr Arg
1               5               10
```

(2) INFORMATION FOR SEQ ID NO:68:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

```
Gly Thr Val Glu Pro Val Pro Pro Pro Val Pro Pro Arg Arg Arg Pro
1               5               10              15

Glu Ser Ala
```

(2) INFORMATION FOR SEQ ID NO:69:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```
Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro Arg Pro Leu Pro
1               5               10              15

Val
```

(2) INFORMATION FOR SEQ ID NO:70:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```
Leu Ala Pro Pro Lys Pro Pro Leu Pro Glu Gly Glu Val
1               5               10
```

(2) INFORMATION FOR SEQ ID NO:71:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

```
     (ix)  FEATURE:
           (A)  NAME/KEY: Modified-site
           (B)  LOCATION: 2
           (D)  OTHER INFORMATION: /note= "any residue other than Cys"

     (ix)  FEATURE:
           (A)  NAME/KEY: Modified-site
           (B)  LOCATION: 5..6
           (D)  OTHER INFORMATION: /note= "any hydrophobic residue"

     (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:71:

     Arg Xaa Leu Pro Xaa Xaa Pro
     1                5
```

(2) INFORMATION FOR SEQ ID NO:72:

```
      (i)  SEQUENCE CHARACTERISTICS:
           (A)  LENGTH: 13 amino acids
           (B)  TYPE: amino acid
           (D)  TOPOLOGY: unknown

     (ii)  MOLECULE TYPE: peptide

     (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:72:

     Pro Gly Phe Arg Glu Leu Pro Pro Leu Pro Pro Ser Arg
     1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:73:

```
      (i)  SEQUENCE CHARACTERISTICS:
           (A)  LENGTH: 13 amino acids
           (B)  TYPE: amino acid
           (D)  TOPOLOGY: unknown

     (ii)  MOLECULE TYPE: peptide

     (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:73:

     Ala Gln Ser Arg Pro Leu Pro Ile Pro Pro Glu Thr Arg
     1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:74:

```
      (i)  SEQUENCE CHARACTERISTICS:
           (A)  LENGTH: 13 amino acids
           (B)  TYPE: amino acid
           (D)  TOPOLOGY: unknown

     (ii)  MOLECULE TYPE: peptide

     (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:74:

     Pro Pro Asn Ser Pro Leu Pro Pro Leu Pro Thr His Leu
     1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:75:

    (i) SEQUENCE CHARACTERISTICS: .
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Thr Gly Arg Gly Pro Leu Pro Pro Leu Pro Asn Asp Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:76:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Tyr Ser Thr Arg Pro Val Pro Pro Ile Thr Arg Pro Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:77:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Ser His Lys Ser Arg Leu Pro Pro Leu Pro Thr Arg Pro
1   5   10

(2) INFORMATION FOR SEQ ID NO:78:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Tyr Arg Phe Arg Ala Leu Pro Ser Pro Pro Ser Ala Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:79:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

    Leu Ala Gln Arg Gln Leu Pro Pro Thr Pro Gly Arg Asp
    1   5   10

(2) INFORMATION FOR SEQ ID NO:80:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

    Ala Leu Gln Arg Arg Leu Pro Arg Thr Pro Pro Pro Ala
    1   5   10

(2) INFORMATION FOR SEQ ID NO:81:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

    Pro Ala Thr Arg Pro Leu Pro Thr Arg Pro Ser Arg Thr
    1   5   10

(2) INFORMATION FOR SEQ ID NO:82:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

    Tyr Ser Thr Arg Pro Leu Pro Ser Arg Pro Ser Arg Thr
    1   5   10

(2) INFORMATION FOR SEQ ID NO:83:

    (i) SEQUENCE CHARACTERISTICS: .
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Xaa Pro Gly Arg Ile Leu Leu Leu Pro Ser Glu Pro Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:84:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Ser Gly Gly Ile Leu Ala Pro Pro Val Pro Pro Arg Asn
1   5   10

(2) INFORMATION FOR SEQ ID NO:85:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Arg Ser Thr Arg Pro Leu Pro Ile Leu Pro Arg Thr Thr
1   5   10

(2) INFORMATION FOR SEQ ID NO:86:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Ser Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:87:

    (i) SEQUENCE CHARACTERISTICS: ·
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Ser Thr Asn Arg Pro Leu Pro Met Ile Pro Thr Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:88:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Arg Ser Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr
1   5   10

(2) INFORMATION FOR SEQ ID NO:89:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Ser Thr Ser Arg Pro Leu Pro Ser Leu Pro Thr Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:90:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Arg Ser Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr
1   5   10

- 86 -

(2) INFORMATION FOR SEQ ID NO:91:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

```
Arg Ser Thr Arg Gln Leu Pro Ile Pro Pro Thr Thr Thr
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:92:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

```
Ser Thr Pro Arg Pro Leu Pro Leu Ile Pro Thr Thr Pro
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:93:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

```
Arg Ser Thr Arg Pro Leu Pro Pro Thr Pro Leu Thr Thr
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:94:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

```
Arg Ser Thr Arg Pro Gln Pro Pro Pro Pro Ile Thr Thr
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:95:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 7 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

  Arg Pro Leu Pro Met Leu Pro
  1   5

(2) INFORMATION FOR SEQ ID NO:96:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 7 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 5
       (D) OTHER INFORMATION: /note= "Pro or Met"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

  Arg Pro Leu Pro Xaa Leu Pro
  1   5

(2) INFORMATION FOR SEQ ID NO:97:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 4 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

  Arg Ser Thr Pro
  1

(2) INFORMATION FOR SEQ ID NO:98:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 14 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

    Arg Ser Thr Pro Ala Pro Pro Val· Pro Pro Arg Thr Thr Arg
    1   5   10

(2) INFORMATION FOR SEQ ID NO:99:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

AGCGTAACGA TCTCCCG   17

(2) INFORMATION FOR SEQ ID NO:100:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TTCACCTCGA AAGCAAGCTG   20

(2) INFORMATION FOR SEQ ID NO:101:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CCTCATAGTT AGCGTAACG   19

(2) INFORMATION FOR SEQ ID NO:102:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

                              - 89 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

AGCGTAACGA TCTAAA   16

(2) INFORMATION FOR SEQ ID NO:103:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

    Lys Ser Gly Ser Gly
    1   5

(2) INFORMATION FOR SEQ ID NO:104:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

    Gly Ser Gly Ser
    1

(2) INFORMATION FOR SEQ ID NO:105:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 11 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

    Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
    1   5   10

(2) INFORMATION FOR SEQ ID NO:106:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

Ser Ser Cys Asp His Thr Leu Gly Leu Gly Trp Cys Gly Ser Arg Ser
1   5   10   15

Thr Arg Gln Leu Pro Ile Pro Pro Thr Thr Thr Arg Pro Ser Arg
            20   25   30

WHAT IS CLAIMED IS:

1.   A peptide having at least nine and up to forty-five
amino acid residues, including an amino acid sequence of the
formula, R-2-L-P-5-6-P-8-9 (SEQ ID NO:10), positioned anywhere
along the peptide, in which each number represents an amino acid
5   residue, such that 2 represents any amino acid residue except
cysteine, 5 and 6 each represents a hydrophobic amino acid
residue, 8 represents any amino acid residue except cysteine, and
9 represents a hydrophilic amino acid residue except cysteine,
each letter being the standard one-letter symbol for the
10   corresponding amino acid, said peptide exhibiting a binding
affinity for the SH3 domain of Src, provided that said peptide is
not R-P-L-P-P-L-P-T-S (SEQ ID NO:11).

2.   The peptide of claim 1 in which 2 is a P, R, A, L, Q, E
or S.

3.   The peptide of claim 1 in which 5 is a P, M, I or L.

4.   The peptide of claim 1 in which 6 is a P, L, I or V.

5.   The peptide of claim 1 in which 8 is a T, R, P, I, N,
E, V, S, A, G or L.

6.   The peptide of claim 1 in which 9 is a T, R, S, H or D.

7.   The peptide of claim 1 which further comprises a
C-terminal-flanking amino acid sequence of the formula 10, 10-11,
10-11-12, 10-11-12-13 (SEQ ID NO:12) or 10-11-12-13-14 (SEQ ID
NO:13), in which each number represents any amino acid residue
5   except cysteine, such that 10 is bound to 9 by a peptide bond.

8.   The peptide of claim 7 in which 10 is T, R, L, S, D, P,
A or N.

9.   The peptide of claim 7 in which 11 is R, P, A, Q, S or
T.

10.   The peptide of claim 7 in which 12 is P, S, R or T.

11.   The peptide of claim 7 in which 13 is P, S, R, F, H or
T.

12.   The peptide of claim 7 in which 14 is S, R, G or T.

13.   The peptide of claim 1 which further comprises an N-
terminal-flanking amino acid sequence of the formula 1', 2'-1',

- 92 -

3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in which each number represents any amino acid residue except cysteine, such that 1' is bound to R by a peptide bond.

14. The peptide of claim 13 in which 1' is T, P, S, N, F, W, K, H, Q or G.

15. The peptide of claim 13 in which 2' is S, T, G, P, R, Q, L, A or H.

16. The peptide of claim 13 in which 3' is R, S, P, G, A, V, Y or L.

17. The peptide of claim 13 in which 4' is R, S, V, T, G, L or F.

18. A peptide having at least thirteen and up to forty-five amino acid residues, including an amino acid sequence of the formula, 3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned anywhere along the peptide, in which each number represents an amino acid residue, such that 3', 2', 1', 2, 8, and 10 each represents any amino acid residue except cysteine, 5 and 6 each represents a hydrophobic amino acid residue, and 9 represents a hydrophilic amino acid residue except cysteine, each letter being the standard one-letter symbol for the corresponding amino acid, said peptide exhibiting a binding affinity for the SH3 domain of Src.

19. The peptide of claim 18 in which 5 is a P or M.

20. The peptide of claim 18 in which 1' is T, P, S or N.

21. The peptide of claim 18 in which 2' is S or T.

22. The peptide of claim 18 in which 3' is R or S.

23. The peptide of claim 18 in which 10 is T or R.

24. The peptide of claim 1 the binding affinity of which is at least three-fold greater than that exhibited by the peptide RPLPPLP for the SH3 domain of Src.

25. The peptide of claim 18 the binding affinity of which is at least three-fold greater than that exhibited by the peptide RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

26. The peptide of claim 1 the binding affinity of which is
at least four-fold greater than that exhibited by the peptide
RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

27. The peptide of claim 18 the binding affinity of which
is at least four-fold greater than that exhibited by the peptide
RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

28. The peptide of claim 1 which further exhibits a general
binding affinity for the SH3 domain of Abl, Grb2, PLC-δ, PLC-γ,
Ras GAP, Nck, p85 PI-3'Kinase, and proteins related thereto.

29. The peptide of claim 1 which exhibits a selective
binding affinity for the SH3 domain of Src and Src-related
proteins.

30. The peptide of claim 18 which further exhibits a
general binding affinity for the SH3 domain of Abl, Grb2, PLC-δ,
PLC-γ, Ras GAP, Nck, p85 PI-3'Kinase, and proteins related
thereto.

31. The peptide of claim 18 which exhibits a selective
binding affinity for the SH3 domain of Src and Src-related
proteins.

32. A peptide having the amino acid sequence RSTPRPLPMLPTTR
(SEQ ID NO:62).

33. A peptide having the amino acid sequence RSTPRPLPPLPTTR
(SEQ ID NO:67).

34. A peptide having the amino acid sequence GILAPPVPPRNTR
(SEQ ID NO:63).

35. A peptide having the amino acid sequence VLKRPLPIPPVTR
(SEQ ID NO:64).

36. A peptide having the amino acid sequence GPHRRLPPTPATR
(SEQ ID NO:65).

37. A peptide having the amino acid sequence ANPSPATRPLPTR
(SEQ ID NO:66).

38. A peptide having an amino acid sequence selected from
the group consisting of RSTRPLPILPRTT, STPRPLPMLPTTR,
STNRPLPMIPTTR, RSTRPLPSLPITT, STSRPLPSLPTTR, RSTRSLPPLPPTT,

- 94 -

RSTRQLPIPPTTT, STPRPLPLIPTTP, RSTRPLPPTPLTT, and RSTRPQPPPPITT

5   (SEQ ID NOS:85-94).

39. A peptide having the amino acid sequence selected from the group consisting of VLKRPLPIPPVTR (SEQ ID NO:64), YSTRPVPPITRPS (SEQ ID NO:76), SHKSRLPPLPTRP (SEQ ID NO:77), GPHRRLPPTPATR (SEQ ID NO:65), PATRPLPTRPSRT (SEQ ID NO:81), and

5   SGGILAPPVPPRN (SEQ ID NO:84).

40. A peptide having the amino acid sequence selected from the group consisting of PPNSPLPPLPTHL (SEQ ID NO:72), TGRGPLPPLPNDS (SEQ ID NO:74), YRFRALPSPPSAS, LAQRQLPPTPGRD ALQRRLPRTPPPA (SEQ ID NOS:78-80), YSTRPLPSRPSRT, and

5   XPGRILLLPSEPR (SEQ ID NOS:82-83).

41. A construct comprising a nucleic acid encoding a peptide of claim 1 or its complement.

42. The construct of claim 41 which is a DNA polynucleotide.

43. The construct of claim 41 which is a RNA polynucleotide.

44. A construct comprising a nucleic acid encoding a peptide of claim 18 or its complement.

45. The construct of claim 44 which is a DNA polynucleotide.

46. The construct of claim 44 which is a RNA polynucleotide.

47. The construct of claim 41 which is a transforming vector.

48. The construct of claim 44 which is a transforming vector.

49. A host cell transformed with the vector of claim 47.

50. A host cell transformed with the vector of claim 48.

51. A conjugate comprising a peptide of claim 1 and a second molecule.

52. The conjugate of claim 51 in which said second molecule is select d from the group consisting of an amino acid, a

- 95 -

peptide, a protein, a nucleic acid, a nucleoside, a glycosidic
residue, a label, a drug or a small molecule.

53. A diagnostic kit for the detection of SH3 domains
comprising an SH3 domain-binding peptide and a detectable label
conjugated to said peptide directly, indirectly or by
complexation, said peptide comprising: (i) a core sequence motif
5   of the formula RXLP$\phi\phi$P (SEQ ID NO:71), wherein X represents any
amino acid except cysteine and ϕ represents a hydrophobic amino
acid residue, each letter representing the standard one-letter
designation for the corresponding amino acid residue; and (ii)
two or more additional amino acid residues flanking said core
10   sequence at its C-terminal end, N-terminal end or both.

54. A drug delivery system comprising an SH3 domain-binding
peptide and a drug conjugated to said peptide directly,
indirectly or by complexation, said peptide comprising: (i) a
core sequence motif of the formula RXLP$\phi\phi$P (SEQ ID NO:71),
5   wherein X represents any amino acid except cysteine and ϕ
represents a hydrophobic amino acid residue, each letter
representing the standard one-letter designation for the
corresponding amino acid residue; and (ii) two or more additional
amino acid residues flanking said core sequence at its C-terminal
10   end, N-terminal end or both.

55. The drug delivery system of claim 54 which may be
administered parenterally, orally, enterally, topically or by
inhalation.

56. The drug delivery system of claim 54 which may be
administered intranasally, opthalmically or intravaginally.

57. The drug delivery system of claim 54 which is in the
form of a solid, gel, liquid or aerosol.

58. A method of modulating the activity of Src or Src-
related proteins comprising administering a composition
comprising an effective amount of a peptide of claim 1 and a
carrier.

59. The method of claim 58 which inhibits the activity of
Src or Src-related proteins.

- 96 -

60. The method of claim 58 which activates Src or Src-related proteins.

61. A method of identifying a peptide having a region that binds to an SH3 domain comprising:

(a) providing an immobilized target protein comprising an SH3 domain;

5
(b) incubating said immobilized target protein with an aliquot taken from a random peptide library;

(c) washing unbound peptide from said immobilized target protein;

(d) recovering the peptide bound to said immobilized

10
target protein; and

(e) determining the primary sequence of the SH3 domain-binding peptide.

62. The method of claim 61 in which said library is a displayed random peptide library.

63. The method of claim 62 in which said library is a phage-displayed random peptide library.

64. The method of claim 62 in which said library is a phagemid-displayed random peptide library.

65. The method of claim 61 in which step (c) includes washing unbound phage from said immobilized target protein; step (d) includes recovering the phage bound to said immobilized target protein; and step (e) includes determining the relevant

5
nucleotide sequence of said binding phage nucleic acid, from which the primary sequence corresponding to the SH3 domain-binding peptide is deduced.

66. A method of identifying a peptide having a region that binds to an SH3 domain comprising:

(a) providing an immobilized target protein comprising an SH3 domain;

5
(b) incubating said immobilized target protein with an aliquot taken from a phage-displayed random peptide library, which library includes peptides having a random sequence of ≥8 amino acid residues;

- 97 -

   (c)   washing unbound phage from said immobilized target
10 protein;

   (d)   recovering the phage bound to said immobilized
target protein; and

   (e)   determining the relevant nucleotide sequence of
said binding phage nucleic acid and deducing the primary sequence
15 corresponding to the SH3 domain-binding peptide.

   67. The method of claim 66 which further comprises
amplifying the titer of the recovered phage.

   68. The method of claim 66 which further comprises
repeating the steps of incubation, washing and recovery to
provide SH3 domain-binding peptide-enriched phage.

   69. A pharmaceutical composition comprising an SH3 domain-
binding peptide and a pharmaceutically acceptable carrier, said
peptide comprising: (i) a 9-mer sequence motif of the formula
RXLP$\phi\phi$PXψ (SEQ ID NO:10), wherein X represents any amino acid
5 except cysteine, ϕ represents a hydrophobic amino acid residue,
and wherein ψ is a hydrophilic amino acid residue except
cysteine, each letter representing the standard one-letter
designation for the corresponding amino acid residue; and,
optionally, (ii) additional amino acid residues flanking said 9-
10 mer sequence at its C-terminal end, N-terminal end or both, up to
a total of 45 amino acid residues, including said 9-mer sequence.

   70. The composition of claim 69 in which at least one
additional amino acid flanks said 9-mer sequence.

   71. The composition of claim 69 in which at least two
additional amino acids flank said 9-mer sequence.

   72. The composition of claim 69 in which at least three
additional amino acids flank said 9-mer sequence.

   73. A method of disrupting protein tyrosine kinase-mediated
signal transduction pathways comprising administering an
effective amount of a peptide of claim 1.

   74. A method of regulating the processing, trafficking or
translation of RNA by administering an effective amount of a
peptide of claim 1.

Fig. 1

```
                            Pro.Gly
        GGC.TCG.AGN.(NNB)₁₈.CCA.GGT
```

$$GGC.TCG.AGN.(NNB)_{18}.CCA.GGT$$

$$GGT.CCA.(NNV)_{18}.AGA.TCT.GG$$

N=A,C,G,T
B=C,G,T
V=A,C,G

↓   Fill in with *Taq* DNA polymerase

Xho I

$$GGC.TCG.AGN.(NNB)_{18}.CCA.GGT$$

$$GGT.CCA.(NNV)_{18}.\underline{AGA.TCT.GG}$$
$$Xba\ I$$

↓   cleave with *Xho* I and *Xba* I

$$TCG.AGN.(NNB)_{18}.CCA.GGT$$

$$GGT.CCA.(NNV)_{18}.AGA.TC$$

↓   ligate with *Xho*I+*Xba*I-cleaved M13 m663 vector

    electroporate into *E. coli* DH5αF'

TSAR-9 library
of pIII-random sequence fusion proteins

$$...S\ H\ S\ \boxed{S\ (R/S)\ X_{18}\ P\ G\ X_{18}\ S\ R}\ P\ S\ R\ T\ ...$$

↑
signal peptidase cleavage site

Fig. 2

$$\text{Gly}$$
$$\underline{\texttt{tt.ttg.tcg.acN.(NNB)}_{10}\texttt{.Ngc.ggt.g}}$$

$$\underline{\texttt{cg.cca.cNV.(NNV)}_{10}\texttt{.tga.tca.ttt.t}}$$

N = A, G, T, C
B = G, T, C
V = G, A, C

↓   Fill in with *Taq* DNA polymerase

Sal I
$$\underline{\texttt{tt.ttg.tcg.acN.(NNB)}_{10}\texttt{.Ngc.ggt.g}}$$

$$\underline{\texttt{cg.cca.cNV.(NNV)}_{10}\texttt{.tga.tca.ttt.t}}$$
Spe I

↓   cleave with *Sal* I+*Spe* I

$$\underline{\texttt{tcg.acN.(NNB)}_{10}\texttt{.Ngc.ggt.g}}$$

$$\underline{\texttt{cg.cca.cNV.(NNV)}_{10}\texttt{.tga.tc}}$$

↓   ligate with *Xho*I+*Xba*I-cleaved M13 m663 vector

      electroporate into *E. coli* DH5αF'

TSAR-12 library
of pIII-random sequence fusion proteins

...S H S |S (S/T) X$_{10}$ Ø G ∂ X$_{10}$ T R| P S R T ...

↑
signal peptidase cleavage site
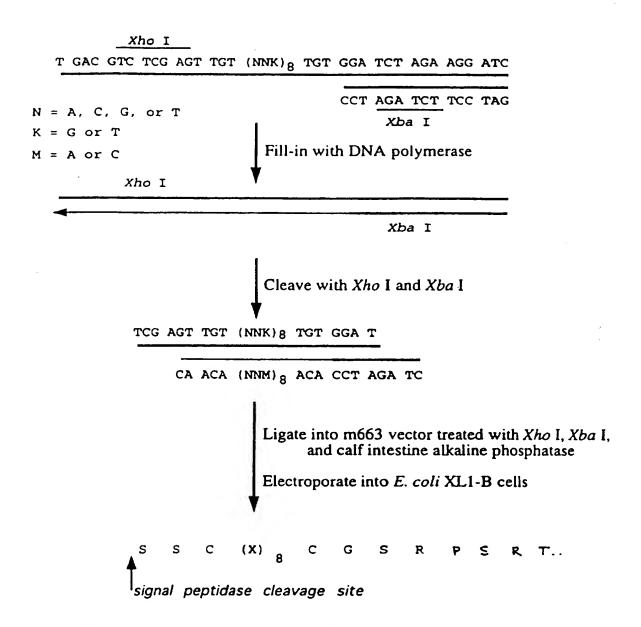
Ø = S, R, G, C, or W

∂ = V, A, D, E, or G

Fig. 3

Construction of R8C library

```
         Xho I
T GAC GTC TCG AGT TGT (NNK)8 TGT GGA TCT AGA AGG ATC

                              CCT AGA TCT TCC TAG
N = A, C, G, or T                    Xba I
K = G or T
M = A or C              │  Fill-in with DNA polymerase
                        ▼
      Xho I
◄─────────────────────────────────────────
                              Xba I
```

```
                    │ Cleave with Xho I and Xba I
                    ▼
TCG AGT TGT (NNK)8 TGT GGA T

CA ACA (NNM)8 ACA CCT AGA TC
```

```
    │  Ligate into m663 vector treated with Xho I, Xba I,
    │     and calf intestine alkaline phosphatase
    │
    │  Electroporate into E. coli XL1-B cells
    ▼
```

```
    S    S    C   (X)8   C    G    S    R    P    S    R    T..
   ▲
   │signal  peptidase  cleavage  site
```

Resultant amino acid sequence expressed at the amino terminus of pIII

3 / 13

Fig. 4

Origin of One Class of Double-insert R8C Recombinants

Xba I

TCG AGT TGT (NNK)$_8$ TGT GGA T CT AGA TCC ACA (MNN)$_8$ ACA AC

CA ACA (NNM)$_8$ ACA CCT A GA TCT AGG TGT (KNN)$_8$ TGT TGA GCT

Xho I   Xho I

Ligation of two double-stranded
oligonucleotides into a head-to-head
arrangement at the Xba I site

Insertion into m663 vector cleaved
by Xho I and Xba I. Illegitimate ligation
at the Xba I site of the vector.

Electroporate into E. coli XL1-B cells

S   S   C (X)$_8$ C G S R S T (X)$_8$ T T R . . . .

signal peptidase cleavage site

Resultant amino acid sequence expressed at the amino terminus of pIII

4 / 13

Fig. 5
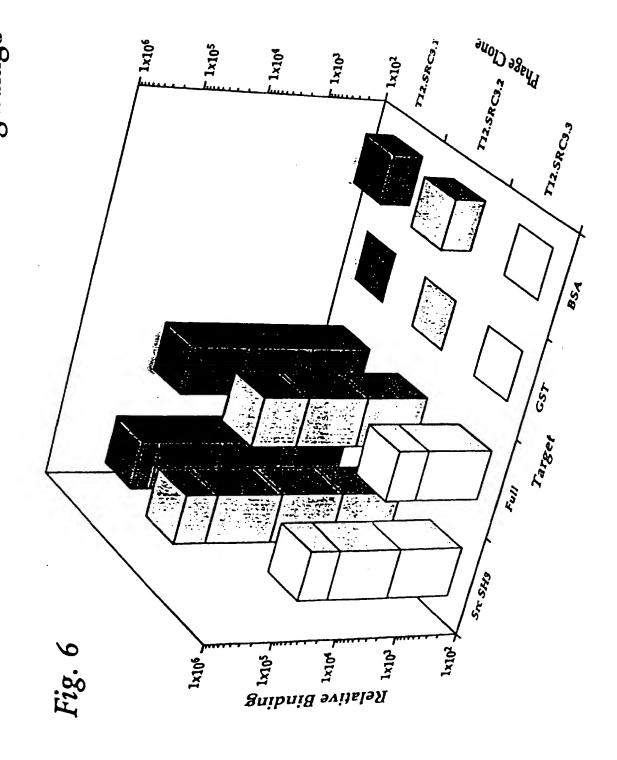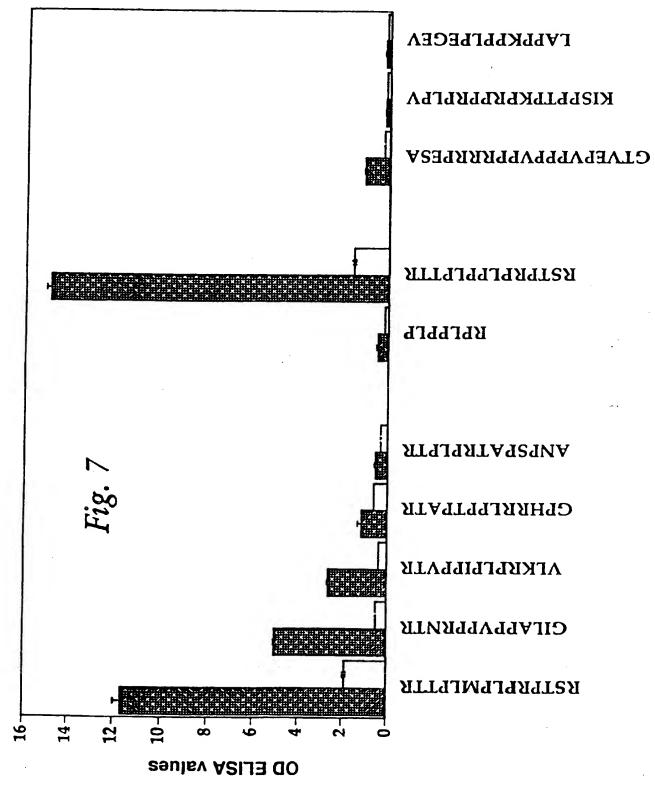
SH3 BINDING PEPTIDES FROM RANDOM PEPTIDE LIBRARIES

CLONE	SEQUENCE			FREQUENCY
T9.SRC3.2	SSFDQQDWDYSIAEKMHPIRPGF	RELPPLP	PSRASFGGGASRPSR	2
T12.SRC3.4	STNVWVTGSVIARGAQS	RPLPIPP	ETRPSR	1
T12.SRC3.6	STAPWGLRVAHEGGVLK	RPLPIPP	VTRPSR	1
T9.SRC3.4	SSSGYVVPKRLGDMREYNAHPGLHVPPN	SPLPPLP	THLQSSRPSR	1
T9.SRC3.6	SSRGEGNNIISSRPFLSNSDRGVSNKLTGR	GPLPPLP	NDSRPSR	2
T12.SRC3.7	STAVSFRFMPGGGGAFYST	RPVPPIT	RPSRT	1
T12.SRC3.5	STAHSLWDGTFSGVSHKS	RLPPLPT	RPSRT	1
T9.SRC3.7	XPGYARIVSYRF	RALPSPP	SASRPSR	1
T12.SRC3.3	STNDVDWMHMWNSGGPH	RRLPPTP	ATRPSR	4
T9.SRC3.5	SSDNWARRVHASELIYTDLSPGILLAQ	RQLPPTP	GRDPSHSRPSR	1
T9.SRC3.1	SSESPLMYNRVGALQSLTSVPGSMMHFALQ	RRLPRTP	PPASRPSR	5
T12.SRC3.2	STRWSHSWPGYVGGANPSPAT	RPLPTRP	SRTVESC	19
T9.SRC3.3	SRYNDLGTRPVSEVIKYDYFPGYSQHVITPDGSYST	RPLPSRP	SRTVESC	2
T9.SRC3.8	XPG	RILLLPS	EPRTFYNYGHDSRPSR	1
T12.SRC3.1	STMYGVSWLSSGSGGILA	PPVPPRN	TRPSR	25
	Consensus	RPLPPLP		
R8C.YES3.6	SSCTEKTVSGWCGSRST	RPLPILP	RTRPSR	1
R8C.YES3.5	SSCMLPTDGWQCGSRSTP	RPLPMLP	TTRPSR	1
R8C.YES3.3/SRC3.1	SSCDTQFRLNCGSRSTN	RPLPMIP	TTRPSR	3
R8C.YES3.1/SRC3.2	SSCMQGQAGLKCGSRST	RPLPSLP	ITRPSR	7
R8C.YES3.7	SSCYREKDTWGCGSRSTS	RPLPSLP	TTRPSR	2
R8C.YES3.2	SSCLFEQGAGTCGSRST	RSLPPLP	PTRPSR	2
R8C.YES3.4	SSCDHTLGLGWCGSRST	RQLPIPP	TTRPSR	1
R8C.YES3.10	SSCDTGRIAPGCGSRSTP	RPLPLIP	TTPRSTNLNLTSTTRPSR	2
R8C.YES3.8	SSCGLDNAAKTCGSRST	RPLPPTP	LTRPSR	2
R8C.YES3.9	SSCSRAHETEMCGSRST	RPQPPPP	ITRPSR	2
	Consensus	RPLPPLP		

Binding Characteristics of Src SH3-Binding Phage

Fig. 6

Fig. 7

Fig. 8

Fig. 9

**Competiti n of the Binding of Cell Lysate Proteins
to Immobilized SH3 Domains**

	peptides: (micromolar concentration)	
12.SEC3.1	GILAPPVPPRNTR	
12.SEC3.6	VLKRPLPIPPVTR	
long consensus	RST P RPLPPLPTTR	
P13 Kinase	KISPPTPKPRPPRPLPV	
Vinculin	LAPPKPPLPEGEV	
12.SEC3.1	GILAPPVPPRNTR	
12.SEC3.6	VLKRPLPIPPVTR	
long consensus	RSTPPRPLPPLPTTR	
P13 Ki...	KKISPPTPKPRPPRPLPV	

Fig. 10

Fig. 11
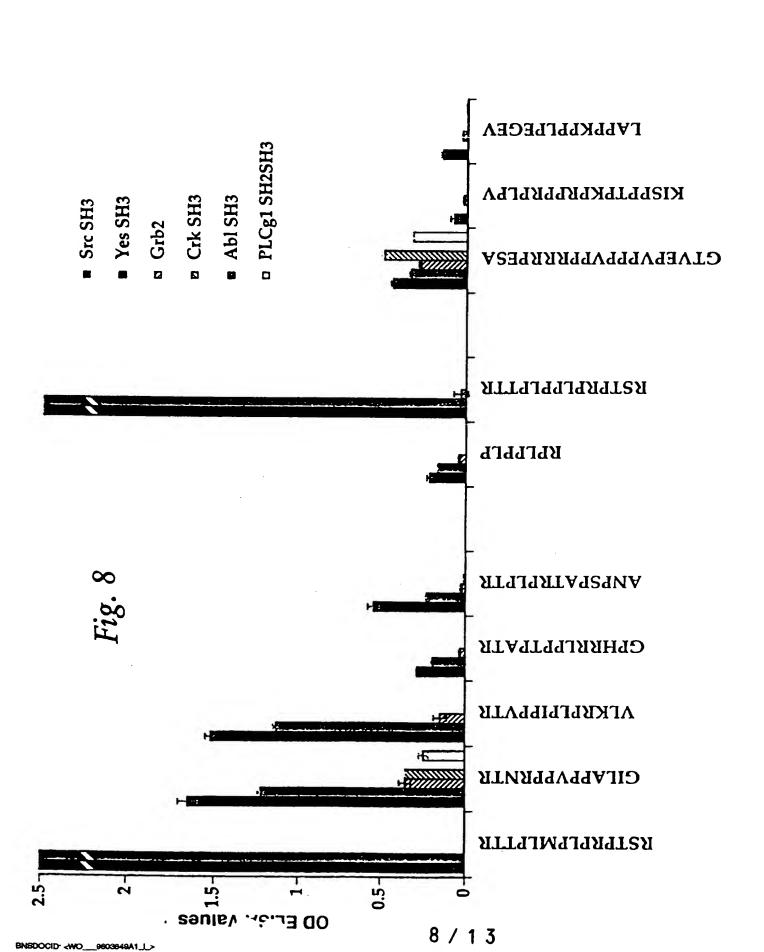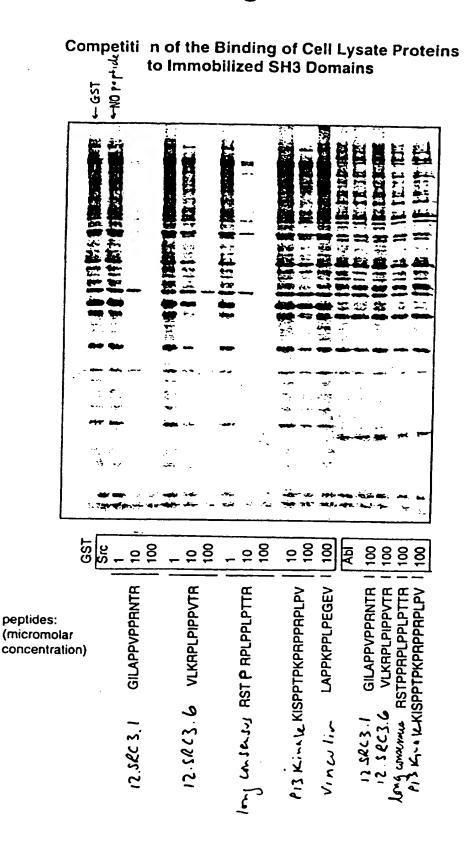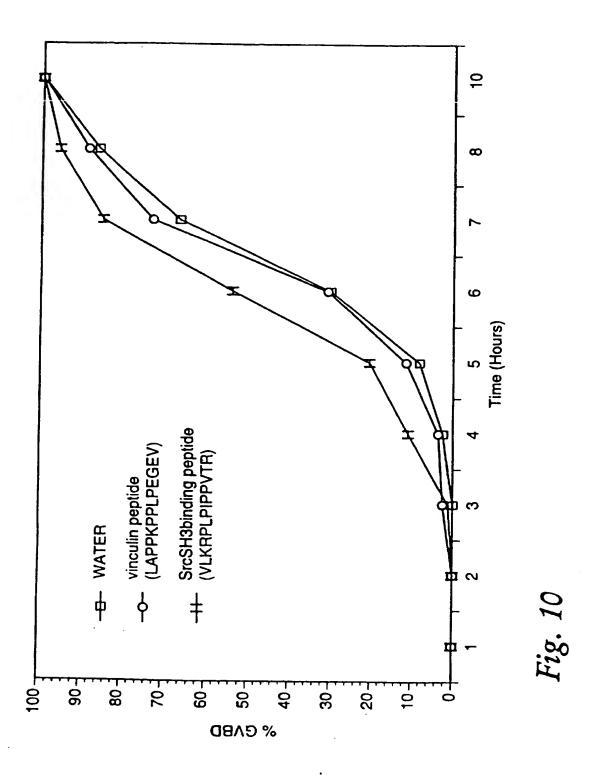
FIGURE 12

ANNEAL OLIGONUCLEOT →
FIL →
DIGEST WI... XhoI AND Xba →
GEL PURIFY CUT INSERT →
CLONE INTO mBAX VECTOr →
EXPRESS B! →

```
CTG TGC CTC GAG K NNK6 CCA NNK2 CCA NNK6 TCT AGA CGT GTC AGT
                                          AGA TCT GCA CAG TCA

CTG TGC CTC GAG K NNK6 CCA NNK2 CCA NNK6 TCT AGA CGT GTC AGT
GAC ACG GAG CTC M NNM6 GGT NNM2     AGA TCT GCA CAG TCA

CTG TGC C   TC GAG K NNK6 CCA NNK2 CCA NNK6 T   CT AGA CGT GTC AGT
GAC ACG GAG CT     C M NNM6 GGT NNM2 GGT NNM6 AGA TC    T GCA CAG TCA

              TC GAG K NNK6 CCA NNK2 CCA NNK6 T
                     C M NNM6 GGT NNM2 GGT NNM6 AGA TC

        ...TC GAG K NNK6 CCA NNK2 CCA NNK6 TCT AGA...
        ...AG CTC M NNMG GGT NNM2 GGT NNM6 AGA CTC...

        S S/R  X6P   X2   P   X6   S   R ..............pIII
```

↓ PEPTIDE CLEAVAGE SITE

FIGURE 13

mBAX

signal peptide
cleavage site

-1 +1
S S | I D M P * T A S T M Y N M L H R N E P |
..TCCTCGAGTATCGACATGCCTTAGACTGCTAGCACTATGTACAACATGCTTCATGGCAACGAGCCA

Xho I

...epitope, mAb...

| G G R K L S P P A N D M P P A L L K R | S R
GGTGGGAGGAAGTTGAGCCCGCCAACGACATGCCGCCCGCCCTCCTGAAGAGAGTCTAGA..

Xba I

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6)   :Please See Extra Sheet.
US CL   :435/5, 7.1, 172.1, 320.1; 530/300; 514/2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. :   435/5, 7.1, 172.1, 320.1; 530/300; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG, STN
search terms: sh3, src, peptides

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	The Journal of Biological Chemistry, Volume 269, Number 39, issued 30 Spetember 1994, Sparks et al., "Identification and Chracterization of Src SH3 Ligands from Phage-displayed Random Peptide Libraries", pages 23853-23856, see entire article.	1-74
X --- Y	Journal of the American Chemical Society, Volume 115, issued 1993, Chen et al. "Biased Combinatorial Libraries: Novel Ligands for the SH3 Domain of Phosphatidyllinositol 3-Kinase", pages 12591-12592, see entire article, especially Tables I and II.	1-6, 24, 26, 28, 29, 53 -------------- 7-23, 25, 27, 30-51, 53-74

X	Further documents are listed in the continuation of Box C.		See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 OCTOBER 1995	0 1 NOV 1995

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	LORA M. GREEN
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Science, Volume 249, issued 28 September 1990, Langer, "New Methods of Drug Discovery", pages 1527-1533, see entire article.	51, 52, 54-57
Y	Science, Volume 257, issued 07 August 1992, Cicchetti et al., "Identificat ion of a Protein That Binds to the SH3 Region of Abl and Is Similar to BCR and GAP-rho"; pages 803-806, see entire article.	7-23, 25, 27, 30-50
Y	Gene, Volume 128, issued 1993, Kay et al., "An M13 Phage Library Displaying Random 38-amino-acid Petides as a Source of Novel Sequences with Affinity to Selected Targets", pages 59-65, see entire article.	7-23, 25, 27, 30-48, 61-68

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

G01N 33/53; C12Q 1/70; C12N 5/10. 7/01; C07K 7/00, 14/00; A61K 38/03

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPI
International B

INTERNATIONAL APPLICATION PUBLISHED UNDER

WO   9603649A1

(51) International Patent Classification ^{6} : G01N 33/53, C12Q 1/70, C12N 5/10, 7/01, C07K 7/00, 14/00, A61K 38/03	**A1**	(11) International Publication Number: **WO 96/03649** (43) International Publication Date: 8 February 1996 (08.02.96)

(21) International Application Number: PCT/US95/09382

(22) International Filing Date: 24 July 1995 (24.07.95)

(30) Priority Data:
278,865   22 July 1994 (22.07.94)   US
483,555   7 June 1995 (07.06.95)   US

(71) Applicant: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; Office of Technology Development, CB #4100, 302 Bynum Hall, Chapel Hill, NC 27599-4100 (US).

(72) Inventors: SPARKS, Andrew, B.; 201 Blue Ridge, Carrboro, NC 27510 (US). KAY, Brian, K.; 18 Wysteria Way, Chapel Hill, NC 27514 (US). THORN, Judith, M.; Apartment 18-G, 200 Barnes, Carrboro, NC 27510 (US). QUILLIAM, Lawrence, A.; Apartment 5, 206 Pinegate Circle, Chapel Hill, NC 27514 (US). DER, Channing, J.; 101 Fieldstone Court, Chapel Hill, NC 27514 (US).

(74) Agents: MISROCK, S., Leslie et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.

(54) Title: Src SH3 BINDING PEPTIDES AND METHODS OF ISOLATING AND USING SAME

(57) **Abstract**

Peptides having general and specific binding affinities for the Src homology region 3 (SH3) domains of proteins are disclosed in the present invention. In particular, SH3 binding peptides have been isolated from three phage-displayed random peptide libraries which had been screened for isolates that bind to bacterial fusion proteins comprising SH3 and glutathione S-transferase (GST). Preferred peptides are disclosed which comprise a core 7-mer sequence (preferably, a consensus motif) and two or more, preferably at least six, additional amino acid residues flanking the core sequence, for a total length of 9, preferably at least 13, amino acid residues and no more than about 45 amino acid residues. Such peptides manifest preferential binding affinities for certain SH3 domains. The preferred peptides exhibit specific binding affinities for the Src-family of proteins. *In vitro* and *in vivo* results are presented which demonstrate the biochemical activity of such peptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CG	Congo		of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

Src SH3 BINDING PEPTIDES AND
METHODS OF ISOLATING AND USING SAME

1.   Field of the Invention

5   The present invention relates to SH3 binding peptides having a broad range of binding specificities. That is, certain members of the SH3 binding peptides disclosed bind with approximately the same facility with SH3 domains derived from different SH3 domain-containing proteins. Other

10 members, in contrast, bind with a much greater degree of affinity for specific SH3 domains. The SH3 binding peptides are obtained from random peptide libraries that are also phage-displayed. Methods are described of obtaining the phage clones that bind to the SH3 domain targets and of

15 determining their relevant nucleotide sequences and consequent primary amino acid sequence of the binding peptides. The resulting SH3 binding proteins are useful in a number of ways, including, but not limited to, providing a method of modulating signal transduction pathways at the

20 cellular level, of modulating oncogenic protein activity or of providing lead compounds for development of drugs with the ability to modulate broad classes, as well as specific classes, of proteins involved in signal transduction.

25   2.   Background of the Invention
          2.1. Src and the SH3 Domain
               Among a number of proteins involved in eukaryotic cell signaling, there is a common sequence motif called the SH3 domain. It is 50-70 amino acids in length,

30 moderately conserved in primary structure, and can be present from one to several times in a large number of proteins involved in signal transduction and in cytoskeletal proteins.
               The protein pp60c-src represents a family of at least nine non-receptor protein tyrosine kinases (NR-PTKs).

35 Memb rs of this family share an overall structural organization comprising a series of catalytic and non-catalytic domains. In Src, a 14-amino-acid myristylation

- 1 -

SUBSTITUTE SHEET (RULE 26)

signal resides at the extr me amino-terminus, and is followed
by a unique region that is not highly conserved among family
members. Following this r gion are two highly conserved 60-
and 100-amino-acid regions, the Src homology (SH) domains 3

5 and 2, respectively. SH2 and SH3 domains have been shown to
play an important role in mediating protein-protein
interactions in a variety of signaling pathways. Koch, C.A.,
et al., in <u>Science</u> (1991) 252:668-74. The carboxy-terminal
half of Src contains the PTK catalytic domain, as well as a

10 negative regulatory tyrosine (Y527) near the carboxy
terminus. Phosphorylation of this residue (e.g., by Csk)
results in the inhibition of PTK activity. Cooper, J.A., et
al., in <u>Science</u> (1986) 231:1431-1434. Mutation of Y527->F
generates forms of Src with increased PTK and oncogenic

15 activity. Cartwright, C.A., et al., in <u>Cell</u> (1987) 49:83-91;
Kmiecik, T.E., et al., in <u>Cell</u> (1987) 49:65-73; and Piwicna-
Worms, H., et al., in <u>Cell</u> (1987) 75-82.

     The fact that some mutations which result in increased
Src PTK and transforming activity map to the Src SH2 (Seidel-

20 Dugan, C., et al., in <u>Mol. Cell. Biol.</u> (1992) 12:1835-45; and
Hirai, H. and Varmus, H.E. in <u>Mol. Cell. Biol.</u> (1990)
10:1307-1318) and SH3 domains (Seidel-Dugan, C., et al.,
supra; Hirai, H. and Varmus, H.E., *supra*; Superti-Furga, G.,
et al., in <u>Embo. J.</u> (1993) 12:2625-34; and Potts, W.M., et

25 al., in <u>Oncogene Res.</u> (1988) 3:343-355) suggests a negative
regulatory role for these domains. That phosphotyrosine
residues within specific sequence contexts represent high
affinity ligands for SH2 domains suggests a model in which
the SH2 domain participates in Y527-mediated inhibition of

30 PTK activity by binding phosphorylated Y527, thereby locking
the kinase domain in an inactive configuration. Matsuda, M.,
Mayer, B.J., et al., in <u>Science</u> (1990) 248:1537-1539. This
model is supported by the observation that phosphopeptides
corresponding to the carboxy-terminal tail of Src bind

35 active, but not inactive, variants of Src. Rouss 1, R.R., et
al., in <u>Proc. Natl. Acad. Sci. U S A</u> (1991) 88:10696-700; and
Liu, X., et al., in <u>Oncogen</u> (1993) 8:1119-1126.

- 2 -

The mechanism of SH3-mediated inhibition of Src PTK activity remains uncl ar. There is evidence that pY527-mediat d inhibition of Src PTK activity involves the SH3 domain as well as the SH2 domain. Okada, M., Howell, et al.,

5 in J. Biol. Chem. (1993) 268:18070-5; Murphy, S.M., et al., in Mol. Cell. Biol. (1993) 13:5290-300; and Superti-Furga, G., et al., *supra*. Although these effects are thought to be a consequence of SH3-mediated protein-protein interactions, precisely how the Src SH3 domain exerts its negative

10 regulatory effect is unclear. Identification of high affinity ligands for the Src SH3 domain could help resolve these issues.

2.2. Protein Tyrosine Kinases and The Immune Response

15   Src-related tyrosine kinases are expressed in a variety of cell types including those of the immune system (lymphocytes, T cells, B cells, and natural killer cells) and the central nervous system (neural cells, neurons, oligodendrocytes, parts of the cerebellum, and the like).

20 Umemori, H. et al., in Brain Res. Mol. Brain Res. (1992) Dec. 16(3-4):303-310. Their presence in these cells and tissues and their interaction with specific cell surface receptors and immunomodulatory proteins (such as T cell antigen receptor, CD14, CD2, CD4, CD40 or CD45) suggest that these

25 kinases serve an important role in the signalling pathways of not only the central nervous system but of the immune system, as well. See, e.g., Ren, C.L. et al., in J. Exp. Med. (1994) 179(2):673-680 (signal transduction via CD40 involves activation of Lyn kinase); Donovan, J.A. and Koretzky, G.A.,

30 in J. Am. Soc. Nephrol. (1993) 4(4):976-985 (CD45, the immune response, and regulation of Lck and Fyn kinases); and Carmo, A.M. et al., in Eur. J. Immunol. (1993) 23(9):2196-2201 (physical association of the cytoplasmic domain of CD2 with p56lck and p59fyn).

35   For instance, mice with disruptions in their Src-like genes, Hck and Fgr, possess macrophages with impaired phagocytic activity or exhibit a novel immunodeficiency

- 3 -

SUBSTITUTE SHEET (RULE 26)

characterized by an increased susc ptibility to infection
with List ria monocytog nes. Lowell, C.A. et al., in <u>Genes</u>
<u>Dev.</u> (1994) 8(4):387-398. Also, it has been shown that
bacterial lipopolysaccharide (LPS) activates CD14-associated
5 p56lyn, p68hck, and p59c-fgr, while inducing the production
of lymphokines, such as TNF-alpha, IL-1, IL-6, and IL-8.
Inhibition of the protein tyrosine kinases blocks production
of TNF-alpha and IL-1.

10   **2.3. SH3 Binding Peptides**
            As mentioned above, it has long been suspected that
SH3 domains are sites of protein-protein interaction, but it
has been unclear what SH3 domains actually bind. Efforts to
identify ligands for SH3 domains have led to the
15 characterization of a number of SH3-binding proteins,
including 3BP1 and 2 (Ren, R., Mayer, et al., in <u>Science</u>
(1993) 259:1157-61), SOS (Olivier, J.P., et al., in <u>Cell</u>
(1993) 73:179-91; and Rozakis-Adcock, M., et al., in <u>Nature</u>
(1993) 363:83-5), p85 PI-3' Kinase (Xingquan, L., et al., in
20 <u>Mol. Cell. Biol.</u> (1993) 13:5225-5232), dynamin (Gout, I., et
al., in <u>Cell</u> (1993) 75:25-36), AFAP-110 (Flynn, D.C., et al.,
in <u>Mol. Cell. Biol.</u> (1993) 13:7892-7900), and CD42 (Barfod,
E.T., et al., in <u>J. Biol. Chem.</u> (1993) 268:26059-26062).
These proteins tend to possess short, proline-rich stretches
25 of amino acids, some of which have been directly implicated
in SH3 binding. A variety of consensus sequences have been
proposed, although the similarity among proline-rich regions
of different SH3-binding proteins tends to be fairly low.
Also, attempts to build consensus sequences are likely
30 complicated by the incorporation of data from proteins that
bind different SH3 domains.
            Thus, Cicchetti, P., et al., in <u>Science</u> (1992) 257:803-
806, published their work relating to the isolation and
sequencing of two naturally-occurring proteins that could be
35 bound *in vitro* by the SH3 domain of the *abl* oncogene product.
Th se workers found that SH3 domains bind short, proline-rich
regions of such proteins. Subsequently, this sam group

- 4 -

SUBSTITUTE SHEET (RULE 26)

disclosed further results (Ren, R. t al., *supra*) in which the SH3 binding sites of the SH3 binding proteins were localized to "a nine- or ten-amino acid stretch rich in proline residues." A consensus sequence incorporating the

5 features of the SH3 binding sites of four SH3 binding proteins was proposed: XPXXPPP¥XP (SEQ ID NO:1), wherein X indicates a position in the amino acid sequence which is not conserved among the four SH3 binding proteins, P represents proline, and ¥ indicates a hydrophobic amino acid residue,

10 such as P or L.

The screening of complex random peptide libraries has been used to identify peptide epitopes for monoclonal (Scott, J.K. and Smith, G.P. in <u>Science</u> (1990) 249:386-390) and polyclonal (Kay, B.K., et al., in <u>Gene</u> (1993) 128:59-65)

15 antibodies, as well as peptide ligands for a variety of proteins, including streptavidin (Devlin, J.J., et al., in <u>Science</u> (1990) 249:404-406; and Lam, K., et al., in <u>Nature</u> (1991) 354:82-84), the endoplasmic reticulum chaperone BiP (Blond-Elguindi, S., et al., in <u>Cell</u> (1993) 75:717-728), and

20 CaM (Dedman, J.R., et al., in <u>J. Biol. Chem.</u> (1993) 268:23025-23030).

Recently, Chen, J.K. et al., in <u>J. Am. Chem. Soc.</u> (1993) 115:12591-12592, described ligands for the SH3 domain of phosphatidylinositol 3-kinase (PI-3' Kinase) which were

25 isolated from a biased combinatorial library. A "biased" library is to be distinguished from a "random" library in that the amino acid residue at certain positions of the synthetic peptide are fixed, i.e., not allowed to vary in a random fashion. Indeed, as stated by these research workers,

30 screening of a "random" combinatorial library failed to yield suitable ligands for a PI-3' Kinase SH3 domain probe. The binding affinities of these unsuitable ligands was described as weak, >100 μM, based on dissociation constants measured by the Biosensor System (BIAcore).

35

More recently, Yu, et al. (Yu, H., et al., in <u>Cell</u> (1994) 76:933-945) used a "biased" synth tic peptide library

- 5 -

SUBSTITUTE SHEET (RULE 26)

of the form XXXPPXPXX (SEQ ID NO:2), wherein X repres nts any
amino acid other than cysteine, to identify a series of
peptides which bind the Src and PI-3' Kinase SH3 domains.
The bias was accomplished by fixing the proline residues at
5 the specific amino acid positions indicated for the "random"
peptide. As stated previously, without this bias, the
technique disclosed fails to identify any SH3 domain-binding
peptides.

A consensus sequence, based on 13 binding peptides was
10 suggested: RXLPPRPXX (SEQ ID NO:3), where X tends to be a
basic residue (like R, K or H). The binding affinities of
several SH3 binding peptides were disclosed as ranging from
8.7 to 30 μM. A "composite" peptide, RKLPPRPRR (SEQ ID
NO:4), was reported to have a binding affinity of 7.6 μM.
15 This value compares favorably to the binding affinity of the
peptide, VPPPVPPRRR (SEQ ID NO:5), to the N-terminal SH3
domain of Grb2. See, Kraulis, P.J. J. Appl. Crystallogr.
(1991) 24:946. Recognizing the limitations of their
technique, Chen and co-workers, *supra*, stated that their
20 results "illustrate the utility of biased combinatorial
libraries for ligand discovery in systems *where there is some
general knowledge of the ligand-binding characteristics of
the receptor*" (emphasis added).

Yu and co-workers, *supra*, further described an SH3
25 binding site consensus sequence, XpØPpXP (SEQ ID NO:6),
wherein X represents non-conserved residues, Ø represents
hydrophobic residues, P is proline, and p represents residues
that tend to be proline. A consensus motif of RXLPPRPXX (SEQ
ID NO:7), where X represents any amino acid other than
30 cysteine, was proposed for ligands of PI-3' Kinase SH3
domain. A consensus motif of RXLPPLPRϕ (SEQ ID NO:8), where
ϕ represents *hydrophobic* residues, was proposed for ligands
of Src SH3 domain. Still, the dissociation constants
reported for the 9-mer peptides ranged only from about 8-70
35 μM and selectivity between one type of SH3 domain and another
was relatively poor, the K_Ds differing by only about a factor
of four.

- 6 -

SUBSTITUTE SHEET (RULE 26)

Henc , th re remains a need to develop techniques for
th id ntification of Src SH3 binding p ptid s which do not
rely on such "biased" combinatorial peptide libraries that
are limited to a partially predetermined set of amino acid
5 sequences. Indeed, the isolation of SH3 binding peptides
from a "random" peptide library has not been achieved
successfully before now. Furthermore, particular peptides
having much greater binding affinities, whether general or
more selective binding for specific SH3 domains, remain to be
10 identified. Binding peptides specific for particular SH3
domains are useful, for example, in modulating the activity
of a particular SH3 domain-containing protein, while leaving
others bearing an SH3 domain unaffected. Still, the more
promiscuous general binding peptides are useful for the
15 modulation of a broad spectrum of SH3 domain-containing
proteins.

The present invention relates to such SH3 binding
peptides, methods for their identification, and compositions
comprising same. In particular, peptides comprising
20 particular sequences of amino acid residues are disclosed.
which were isolated from random peptide libraries. In the
present invention, clones were isolated from a phage-
displayed random peptide library which exhibited strong
binding affinities for SH3 domain-containing protein targets.
25 Some of these protein targets, include Abl, Src, Grb2, PLC-δ,
PLC-γ, Ras GAP, Nck, and p85 PI-3' Kinase. From the
nucleotide sequence of the binding phage, the amino acid
sequence of the peptide inserts has been deduced. Synthetic
peptides having the desired amino acid sequences are shown to
30 bind the SH3 domain of the target proteins. In particular,
synthetic peptides combining a core consensus sequence and
additional amino acid residues flanking the core sequence are
especially effective at binding to particular target protein
SH3 domains. The SH3 binding peptides disclosed herein can
35 be utilized in a number of ways, including th potential
modulation of oncog nic protein activity *in vivo*. These
peptides also serve as useful leads in the production of

- 7 -

SUBSTITUTE SHEET (RULE 26)

peptidomimetic drugs that modulate a large class of proteins
involved in signal transduction pathways and oncogenesis.

3.   Summary of the Invention

5   Accordingly, three phage-displayed random peptide
libraries were screened for isolates that bind to bacterial
fusion proteins consisting of the Src homology region 3 (SH3)
and glutathione S-transferase (GST). DNA sequencing of the
isolates showed that they contained sequences that resembl
10 the consensus motif, RPLPPLP (SEQ ID NO:9), within their 8,
22, or 36 amino acid long random regions. When peptides were
synthesized corresponding to the pIII inserts of the SH3-
binding phage, they bound to the GST fusions of the SH3
domains of Src and the Src-related proteins, such as Yes, but
15 not of Grb2, Crk, Abl, or PLCγ1. The synthesized peptides
bind quite well to the Src SH3 domain and act as potent
competitors of natural Src SH3 interactions in cell lysates.
For instance, these peptides can compete with radiolabelled
proteins from cell lysates in binding to immobilized Src-GST,
20 with an apparent IC_{50} of 1-10 μM. When a peptide, bearing the
consensus sequence RPLPPLP (SEQ ID NO:9) was injected into
Xenopus laevis oocytes, it accelerated the rate of
progesterone-induced maturation. These results demonstrate
the utility of phage-displayed random peptide libraries in
25 identifying SH3-binding peptide sequences and that such
identified peptides exhibit both *in vivo* and *in vitro*
biological activity.
          Thus, it is an object of the present invention to
provide peptides having at least nine and up to forty-five
30 amino acid residues, including an amino acid sequence of the
formula, R-2-L-P-5-6-P-8-9 (SEQ ID NO:10), positioned
anywhere along the peptide, in which each number represents
an amino acid residue, such that 2 represents any amino acid
residue except cysteine, 5 and 6 each represents a
35 hydrophobic amino acid residue, 8 repres nts any amino acid
residue except cysteine, and 9 repr sents a hydrophilic amino
acid residue except cysteine, each l tter being th standard

- 8 -

one-letter symbol for the corresponding amino acid, said
peptide exhibiting a binding affinity for the SH3 domain of
Src, provid d that said peptid is not R-P-L-P-P-L-P-T-S (SEQ
ID NO:11). In a particular embodiment of the present
5 invention, the peptides also exhibit a binding affinity for
the SH3 domain of Src-related proteins, including Yes, Fyn,
Lyn, Lck, Hck and Fgr.

The present invention also contemplates SH3 domain-
binding peptides that further comprise a C-terminal-flanking
10 amino acid sequence of the formula 10, 10-11, 10-11-12, 10-
11-12-13 (SEQ ID NO:12) or 10-11-12-13-14 (SEQ ID NO:13), in
which each number represents any amino acid residue except
cysteine, such that 10 is bound to 9 by a peptide bond.
Furthermore, peptides are also provided which further
15 comprise an N-terminal-flanking amino acid sequence of the
formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in
which each number represents any amino acid residue except
cysteine, such that 1' is bound to R by a peptide bond.

Thus, in a particular embodiment, a peptide is disclosed
20 having at least thirteen and up to forty-five amino acid
residues, including an amino acid sequence of the formula,
3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned
anywhere along the peptide, in which each number represents
an amino acid residue, such that 3', 2', 1', 2, 8, and 10
25 each represents any amino acid residue except cysteine, 5 and
6 each represents a hydrophobic amino acid residue, and 9
represents a hydrophilic amino acid residue except cysteine,
each letter being the standard one-letter symbol for the
corresponding amino acid, said peptide exhibiting a binding
30 affinity for the SH3 domain of Src.

The present invention also seeks to provide new
consensus sequences or motifs that reflect variations in SH3
domain binding selectivities or specificities. The present
invention also contemplates conjugates of the SH3 binding
35 peptides and a second mol cule or ch mical moiety. This
s cond molecule may be any desired substance whose d liv ry
to the region of the SH3 domain of a particular protein (or

- 9 -

cell containing the protein) is s ught. Possible target
cells include, but are not limited to, neural cells, immune
cells (e.g., T cells, B cells, natural killer c lls, and the
like), osteoclasts, platelets, epidermal cells, and the lik ,
5 which cells express Src, Src-related proteins, and
potentially, other SH3 domain-containing proteins. In this
manner, the modulation of the biological activity of proteins
bearing an SH3 domain can be accomplished.

Other methods and compositions consistent with the
10 objectives of the present invention are likewise disclosed.
In particular, a method is disclosed of modulating the
activity of Src or Src-related proteins comprising
administering a composition comprising an effective amount of
a peptide of the present invention and a carrier, preferably
15 a pharmaceutically acceptable carrier. In a specific
embodiment, the contemplated method results in the inhibition
of the activity of Src or Src-related proteins.
Alternatively, the method is effective to activate Src or
Src-related proteins.

20   In yet another embodiment, a method is disclosed of
identifying a peptide having a region that binds to an SH3
domain comprising: (a) providing an immobilized target
protein comprising an SH3 domain; (b) incubating the
immobilized target protein with an aliquot taken from a
25 random peptide library; (c) washing unbound library peptid s
from the immobilized target protein; (d) recovering the
peptide bound to the immobilized target protein; and (e)
determining the primary sequence of the SH3 domain-binding
peptide.

30   Moreover, a method is disclosed of imaging cells,
tissues, and organs in which Src or Src-related proteins are
expressed, which comprises administering an effective amount
of a composition comprising an SH3 domain-binding peptide
conjugated to detectable label or an imaging agent.

35   Other objectiv s of the present invention will become
apparent to one of ordinary skill in the art after

- 10 -

consideration of the abov disclosure and the following
detailed d scription of the preferred embodiments.

4.   Brief Description of the Figures

5   FIG. 1 illustrates a scheme for the generation of a
random 36 amino acid peptide library (TSAR-9; e.g., SEQ ID
NO:16). Oligonucleotides were synthesized (SEQ ID NOS:17-
18), converted into double-stranded DNA, cleaved with
restriction enzymes (SEQ ID NOS:19-20), and cloned into the
10 M13 vector, m663. The random peptide region encoded by the
oligonucleotides is shown in the box (SEQ ID NO:16) and is
situated at the N-terminus of mature protein III (SEQ ID
NO:21).

        FIG. 2 illustrates a scheme for the generation of a
15 random 22 amino acid peptide library (TSAR-12; e.g., SEQ ID
NO:23). Oligonucleotides were synthesized (SEQ ID NOS:24-
25), converted into double-stranded DNA, cleaved with
restriction enzymes (SEQ ID NOS:26-27), and cloned into the
M13 vector, m663. The random peptide region encoded by the
20 oligonucleotides is shown in the box (SEQ ID NO:23) and is
situated at the N-terminus of mature protein III (SEQ ID
NO:28).

        FIG. 3 illustrates a scheme for the generation of a
random 8 amino acid peptide library (R8C; SEQ ID NO:30).
25 Oligonucleotides were synthesized (SEQ ID NOS:31-32),
converted into double-stranded DNA, cleaved with restriction
enzymes (SEQ ID NOS:33-34), and cloned into the M13 vector,
m663. The random peptide region (SEQ ID NO:30) is flanked by
cysteine residues and is situated at the N-terminus of mature
30 protein III (SEQ ID NO:35).

        FIG. 4 illustrates the possible origin of one class of
double-insert R8C recombinants (e.g., encoding SEQ ID NO:36).
Double-stranded oligonucleotides (e.g., SEQ ID NO:37) may
have ligated in a head-to-head fashion at the *Xba* I site
35 prior to cloning in the *Xho* I- *Xba* I cleaved M13 vector.

        FIG. 5 shows a list of random peptide recombinants (SEQ
ID NOS:38-61 and 106) isolated by the method of the present

- 11 -

SUBSTITUTE SHEET (RULE 26)

invention and the displayed peptide sequence. The amino acid
sequences are aligned to highlight the core sequences. The
flanking s quences are shown to the N-terminal and C-terminal
ends of the core sequence.

5   FIG. 6 graphically illustrates the relative binding
affinities of selected phage clones for various SH3 domains.
The results indicate that certain amino acid sequences
provide generic SH3 domain binding, while others can provide
greater selectivity for the SH3 domain of Src. Still oth r
10 clones exhibit Src SH3 domain preferential binding.

        FIG. 7 shows the binding of synthetic peptides (SEQ ID
NOS:9 and 62-70) representing Src SH3-selected phage inserts
to Src SH3-GST fusion target (shaded columns) over background
GST binding (unshaded columns) relative to the core peptide
15 RPLPPLP (SEQ ID NO:9) and proline-rich peptide segments
derived from naturally occurring proteins. Bound
biotinylated peptide was detected with streptavidin-alkaline
phosphatase ELISA. Each point was performed in triplicate;
average absorbance at 405 nm is presented. Error bars
20 represent SD.

        FIG. 8 illustrates the relative specificity of select d
peptides (SEQ ID NOS:9 and 62-70) for SH3 domains derived
from different proteins. In particular, the binding
affinities of the peptides for the SH3 domains of the
25 following protein fusion targets were tested: Src SH3-GST,
Yes SH3-GST, Grb2-GST, Crk SH3-GST, Abl SH3-GST, PLCγ1
SH2SH3-GST. Bound biotinylated peptide was detected with
streptavidin-alkaline phosphatase. Each point was perform d
in triplicate; values are average signal (absorbance at 405
30 nm) above GST background, with error bars representing
standard deviation. Hatched bars indicate saturation of the
ELISA signal.

        FIG. 9 presents the results of competition experiments
in which selected peptides were found to inhibit the binding
35 of proteins from cell lysates to immobilized Src SH3-GST or
Abl SH3-GST prot in fusion targets.

- 12 -

SUBSTITUTE SHEET (RULE 26)

FIG. 10 pr sents a graph illustrating the increas d rate of progesterone-induced maturation of oocytes injected with an SH3 domain-binding peptide, VLKRPLPIPPVTR (SEQ ID NO:64), of the present invention. Briefly, Stage VI oocyted were -

5 prepared and injected as previously described (see, Kay, B.K., in Methods in Cell Biol. (1991) 36:663-669). Oocytes were injected with 40 nL of 100 μM test peptide or water. After injection, the oocytes were placed in 2 μg/mL progesterone (Sigma, St. Louis, MO) and scored hourly for

10 germinal vesicle breakdown (GVBD).

FIG. 11 shows the results of fluorescence experiments in which certain peptides, Panel A = VLKRPLPIPPVTR (SEQ ID NO:64), Panel B = GILAPPVPPRNTR (SEQ ID NO:63), Panel C = RSTPRPLPPLPTTR (SEQ ID NO:67), of the invention were shown to

15 localize within cellular compartments thought to contain Src or Src-related proteins.

FIG. 12 illustrates a scheme for the generation of a biased peptide library. Oligonucleotides were synthesized, converted into double-stranded DNA, cleaved with restriction

20 enzymes XhoI and XbaI, and cloned into the mBAX vector, described further below in the Examples section. The biased peptide region is situated at the N-terminus of mature pIII protein.

FIG. 13 illustrates the peptide sequence encoded in the

25 mBAX vector situated at the N-terminus of mature pIII protein.

5.   Detailed Description of the Preferred Embodiments
5.1. General Considerations

30   The present invention relates to peptides that exhibit a binding affinity for an SH3 domain, which domain has been found to be present in a number of physiologically significant proteins. In particular, peptides are disclosed which exhibit general binding characteristics to the SH3

35 domains found in a group of proteins, including but not limited to Abl, Src, Grb2, PLC-δ, PLC-γ, Ras GAP, Nck, and p85 PI-3' Kinase. Preferred peptid s xhibit sel ctive, if

- 13 -

not specific, binding affinity for the SH3 domain of Src. As
d scrib d herein, the peptides of the present inv ntion
include a core sequence, preferably a consensus seqeunce, and
additional amino acid residues that flank the core sequence.
5 These peptides, including the methods for their
identification, are described in greater detail, below.

   Thus, in a specific embodiment of the invention,
peptides are provided which have at least nine and up to
about forty-five amino acid residues, including an amino acid
10 sequence resembling the formula,

    R-2-L-P-5-6-P-8-9 (SEQ ID NO:10),

positioned anywhere along the peptide. In the above-
mentioned formula, each number represents an amino acid
residue, such that 2 represents any amino acid residue except
15 cysteine, 5 and 6 each represents a hydrophobic amino acid
residue, 8 represents any amino acid residue except cysteine,
and 9 represents a hydrophilic amino acid residue except
cysteine. Each letter used in the formulas herein represent
the standard one-letter symbol for the corresponding amino
20 acid. When the peptide is a 9-mer, the peptide
R-P-L-P-P-L-P-T-S (SEQ ID NO:11) is excluded. The peptides
of particular interest are those that exhibit a binding
affinity for the SH3 domain of Src and Src-related proteins,
including Yes, Fyn, Lyn, Lck, Hck and Fgr. Preferably, the
25 peptides of the invention exhibit a binding affinity for the
SH3 domain of Src, which is at least three-fold, more
preferably at least four-fold, most preferably at least about
five-fold greater than that exhibited by the peptide RPLPPLP
(SEQ ID NO:9). In still other embodiments, the peptides
30 exhibit a binding affinity for the SH3 domain of Src which is
at least ten-fold greater than that exhibited by the peptid
RPLPPLP (SEQ ID NO:9).

   In specific embodiments, peptides are disclosed in which
the various amino acid residues at the indicated positions
35 may independently have the following preferred identities: 2
is a P, R, A, L, Q, E or S, more preferably P or R; 5
represents a P, M, I or L, more pref rably P or M; 6 is a P,

- 14 -

SUBSTITUTE SHEET (RULE 26)

L, I or V, more preferably P or L; 8 is a T, R, P, I, N, E,
V, S, A, G or L, more preferably T or R; and 9 is a T, R, S,
H or D, more preferably T or R. Despite the preference for
hydrophobic amino acid residues at 5 and 6, in some cases it
5 may be desirable to have hydrophilic amino acid residues at
these positions. Specifically, amino acid residue 5 may be a
T, R or S, and amino acid residue 6 may be a T or R.
Likewise, while a hydrophilic amino acid residue is preferred
at position 9, in some instances a hydrophobic residue, such
10 as a P or A, may be desirable.

The present invention also contemplates SH3 domain-
binding peptides with a minimum length of 10, 11, 12, 13, 14,
15 or more amino acids. Such peptides contain additional
amino acid residues flanking the core sequence of
15 R-2-L-P-5-6-P (SEQ ID NO:71) either at the C-terminal end,
the N-terminal end or both. Thus, for example, such peptides
include those that further comprise a C-terminal-flanking
amino acid sequence of the formula 10, 10-11, 10-11-12, 10-
11-12-13 (SEQ ID NO:12) or 10-11-12-13-14 (SEQ ID NO:13), in
20 which each number represents any amino acid residue except
cysteine, such that the amino acid residue 10 is bound to the
amino acid residue 9 by a peptide bond. In that case,
specific embodiments include an amino acid residue 10 which
is T, R, L, S, D, P, A or N, preferably T or R, an amino acid
25 residue 11 which is R, P, A, Q, S or T, preferably R or P, an
amino acid residue 12 which is P, S, R or T, preferably P or
S, an amino acid residue 13 which is P, S, R, F, H or T,
preferably P or S, and an amino acid residue 14 which is S,
R, G or T, preferably, S or R.

30   Furthermore, peptides are also provided which further
comprise an N-terminal-flanking amino acid sequence of the
formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in
which each number represents any amino acid residue except
cysteine, such that 1' is bound to R by a peptide bond. In
35 such a cas , specific embodiments are provided in which the
amino acid r sidue 1' is T, P, S, N, F, W, K, H, Q or G,
preferably T or P, wherein the amino acid residue 2' is S, T,

- 15 -

SUBSTITUTE SHEET (RULE 26)

G, P, R, Q, L, A or H, preferably S or T, wherein the amino acid residue 3' is R, S, P, G, A, V, Y or L, preferably S or T, and wherein the amino acid residue 4' is R, S, V, T, G, L or F, preferably R or S.

5   In a particular embodiment, a peptide is disclosed having at least thirteen and up to forty-five amino acid residues, including an amino acid sequence of the formula, 3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned anywhere along the peptide, in which each number represents

10 an amino acid residue, such that 3', 2', 1', 2, 8, and 10 each represents any amino acid residue except cysteine, 5 and 6 each represents a hydrophobic amino acid residue, and 9 represents a hydrophilic amino acid residue except cysteine, each letter being the standard one-letter symbol for the

15 corresponding amino acid, said peptide exhibiting a binding affinity for the SH3 domain of Src. Preferred 13-mers include, but are not limited to, those having an amino acid residue 5 which is a P or M, an amino acid residue 1' which is T, P, S or N, an amino acid residue 2' which is S or T, an

20 amino acid residue 3' which is R or S, and an amino acid residue 10 which is T or R. In all the SH3 domain-binding peptides described herein, the prohibition against the use of the hydrophilic amino acid residue cysteine (C) does not extend beyond the 7-mer "core" sequence and the additional

25 amino acid residues flanking the core up to a total (core + flanking) of about 20 amino acids. That is, the occasional use of a cysteine is not absolutely prohibited. What should be kept in mind is that the potential for the formation of intramolecular disulfide bonds, to form a cyclic structure,

30 be minimized as much as possible. Applicants have found that cyclized structures appear to be disfavored, at least with potential binding peptides of less than about 15 amino acid residues in length. The concern for the formation of cyclized structures comprising the core sequence diminishes

35 with increasing size of the peptide. Presumably, a large enough structure, though cyclic, may allow the critical cor sequence to adopt a more or l ss linear conformation.

- 16 -

SUBSTITUTE SHEET (RULE 26)

In particular, specific peptides are disclosed which
exhibit binding affinities to SH3 domains. These include the
peptides, RSTPRPLPMLPTTR (SEQ ID NO. 62), RSTPRPLPPLPTTR (SEQ
ID NO. 67), GILAPPVPPRNTR (SEQ ID NO. 63), VLKRPLPIPPVTR (SEQ
5 ID NO. 64), GPHRRLPPTPATR (SEQ ID NO. 65), and ANPSPATRPLPTR
(SEQ ID NO. 66).

Phage clones are also disclosed, along with the amino
acid sequences that are responsible for SH3 domain binding.
These phage clones are identified in Figure 5.

10   In other embodiments of the present invention, SH3
domain-binding peptides are contemplated which have a total
of 11, 13, 14, 18, 20, 22, 23, 25, 30, 36, 38 or 45 amino
acid residues.

The peptides of the present invention, having been
15 disclosed herein, may be prepared by any number of
practicable methods, including but not limited to solution-
phase synthesis, solid-phase synthesis, protein expression by
a transformed host, cleavage from a naturally-derived,
synthetic or semi-synthetic polypeptide, or a combination of
20 these techniques.

The SH3 binding peptides exhibit a wide range of
biological activity which includes the enhancement (or
inhibition, depending on the particular peptide or the nature
of the peptide's target molecule, in this case a protein
25 bearing an SH3 domain) of the natural function or biological
activity of the peptide's target molecule. For example, the
interaction of the binding peptide of the present invention
could result in the modulation of the oncogenic activity of
the target molecule bearing the SH3 domain. If the target
30 molecule has, in turn, a natural binding partner or ligand,
the peptides of the present invention may also exhibit
antagonistic or agonistic activity in relation to the
biological activity of the natural binding partner.

Thus, it is an object of the present invention to
35 provide a method of activating Src or Src-related protein
tyrosine kinases by administering an effective amount of the
SH3 domain-binding peptides generally described herein. Th

- 17 -

int nsity of th immune response can thus be stimulated, for example, by the incr ased production of certain lymphokines, such as TNF-alpha and interleukin-1. As is gen rally known to those of ordinary skill in the art, a more intense immune

5 response may be in order in certain conditions, such as in combating a particularly tenacious infection, viral or otherwise, or a malignancy.

Furthermore, in a specific embodiment of the present invention, a conjugate compound is contemplated which

10 comprises the peptide of the present invention and a second chemical moiety. The second chemical moiety can be selected from a wide variety of chemical compounds including the peptide itself. Typically, however, the second chemical moiety is selected to be other than the peptide of the

15 present invention, including but not limited to an amino acid, a peptide other than an SH3 binding peptide of the present invention, a polypeptide or protein (i.e., the conjugate is a fusion protein), a nucleic acid, a nucleoside, a glycosidic residue (i.e., any sugar or carbohydrate), a

20 label or image-enhancing agent (including metals, isotopes, radioisotopes, chromophores, fluorophores (such as FITC, TRITC, and the like), and enzyme substrates), a drug (including synthetic, semisynthetic, and naturally-occurring compounds), small molecules (e.g., biotin, hormones, factors)

25 and the like.

The peptide of the present invention can be conjugated to the second chemical moiety either directly (e.g., through appropriate functional groups, such as an amine or carboxylic acid group to form, for example, an amine, imine, amide,

30 ester, acyl or other carbon-carbon bond) or indirectly through the intermediacy of a linker group (e.g., an aliphatic or aromatic polyhydroxy, polyamine, polycarboxylic acid, polyolefin or appropriate combinations thereof). Moreover, the term "conjugate," as used herein, is also meant

35 to nc mpass non-covalent interactions, including but not limited to ionic, affinity or other complexation interactions. Preferably, such other non-covalent

- 18 -

SUBSTITUTE SHEET (RULE 26)

interactions provide definable, most preferably, isolatable chemical conjugate species.

As described further herein, th peptid s of the present invention have been shown to localize within certain cellular

5 compartments which contain Src or Src-related proteins. Consequently, the above-described conjugate can be utilized as a delivery system for introduction of a drug to cells, tissues or organs that include SH3 domain-containing proteins.

10   It should also be pointed out that the present invention seeks to provide a recombinant construct comprising a nucleic acid or its complement that includes codons or nucleotide sequences encoding a peptide having a region that binds to an SH3 domain, preferably the Src SH3 domain. The recombinant

15 nucleic acid may be a DNA or RNA polynucleotide.

In a specific embodiment, the present invention contemplates a recombinant construct which is a transforming vector. Such vectors include those well known to those of ordinary skill in the art, which effect the transfer or

20 expression of the nucleotide sequence after introduction to a host, such as recombinant plasmid, phage or yeast artificial chromosome. These vectors may be closed circular loops or they may be linearized. The vectors contemplated include those that exist extrachromosomally after host transformation

25 or transfection, as well as those that integrate within or even displace portions of the host chromosome. The vectors may be introduced to the cell with the help of transfection aids or techniques well-known in the art. For example, these aids or techniques may take the form of electroporation, use

30 of calcium chloride, calcium phosphate, DEAE dextran, liposomes or polar lipid reagents known as LIPOFECTIN or LIPOFECTAMINE. In addition, the present invention contemplates the direct introduction of the desired nucleic acid to the host cell, for instance, by injection.

35   Transformed host cells are also obtained by the methods of the present invention which are capable of reproducing the polynucleotide s quences of interest and/or expressing the

- 19 -

SUBSTITUTE SHEET (RULE 26)

corresponding peptide products. A variety of hosts are
contemplated, including prokaryotic and eukaryotic hosts. In
particular, bacterial, viral, yeast, animal, and plant cells
are potentially transformable hosts. Thus, a method is

5 disclosed to obtain a transformed host cell that can produce,
preferably secrete, a peptide having a region that binds to
an SH3 domain comprising (a) providing an expression vector,
preferably a secretory expression vector, comprising a
nucleotide sequence encoding at least one copy of a peptide

10 having a region that binds to an SH3 domain; and (b)
introducing the vector to a competent host cell.

The peptides, thus produced, may then be introduced to
cells, tissues, organs, or administered to the subject for
the purpose of modulating the biochemical activity of the SH3

15 domain-containing proteins present therein. Accordingly, in
specific embodiments of the present invention, compositions
are provided which comprise an SH3 domain-binding peptide,
including a core sequence and flanking sequences, and a
suitable carrier.

20   The compositions contemplated by the present invention
may also include other components, from those that facilitate
the introduction or administration of the compositions to
those that have their own innate activity, such as a
prophylactic, a diagnostic or a therapeutic action. Such

25 innate activity may be distinct from that of the peptides of
the present invention or be complementary thereto. In any
event, the compositions of the present invention include
those that are suitable for administration into mammals,
including humans. Preferably, the compositions (including

30 necessarily the carrier) of the present invention are
sterile, though others may need only be cosmetically,
agriculturally or pharmaceutically acceptable. Still other
compositions may be adapted for veterinary use.

The compositions, including the drug delivery systems

35 described herein, are contemplated to be administered in a
vari ty of ways, such as parenterally, orally, nterally,
topically or by inhalation. The compositions may also be

- 20 -

SUBSTITUTE SHEET (RULE 26)

adminster d intranasally, opthalmically or intravaginally.
Furthermor , th compositions of the invention can take
several forms, such as solids, g ls, liquids, aerosols or
patches.

5   In another embodiment of the present invention a method
is provided of identifying a peptide having a region that
binds to an SH3 domain comprising: (a) providing an
immobilized target protein comprising an SH3 domain; (b)
incubating the immobilized target protein with an aliquot
10 taken from a phage-displayed random peptide library, which
library includes peptides having a random sequence of ≥8
amino acid residues; (c) washing unbound phage from the
immobilized target protein; (d) recovering the phage bound to
the immobilized target protein; and (e) determining the
15 relevant nucleotide sequence of said binding phage nucleic
acid and deducing the primary sequence corresponding to the
SH3 domain-binding peptide. Preferably, the method further
comprises amplifying the titer of the recovered phage and
repeating the steps of incubation, washing and recovery to
20 provide SH3 domain-binding peptide-enriched phage.

Any other mode by which the peptide library, random or
otherwise, can be "displayed" can be utilized in the present
invention, however. Moreover, the present applicants believe
that longer *random* peptide sequences (e.g., >6 amino acid
25 residues, preferably >10, and most preferably, >12) provide
not only much greater diversity but also a richer degree of
secondary structure conducive to binding activity. If the
random region of the peptide is less than or equal to an 8-
mer, it should preferably not be cyclized.

30

5.2. Preparation of Random Peptide Libraries

The preparation and characterization of the
preferred phage-displayed random peptide libraries have been
described elsewhere. See, for example, Kay, B.K. et al. in
35 <u>Gene</u> (1992) 128:59-65, for a description of the preparation
of the phage-displayed random peptide library known as TSAR-
9, more b low. In particular, by cloning degenerate

- 21 -

oligonucleotides of fixed length into bacteriophag vectors, recombinant librari s of random peptides can be generated which are expressed at the amino-terminus of the pIII protein on the surface of M13 viral particles. (There are 3-5 copies
5 of the pIII-fusion on the surface of each particle.) Phag display offers several conveniences: first, the expressed peptides are on the surface of the viral particles and accessible for interactions; second, the recombinant viral particles are stable (i.e., can be frozen, exposed to pH
10 extremes); third, the viruses can be amplified; and fourth, each viral particle contains the DNA encoding the recombinant genome. Consequently, these libraries can be screened by isolating viral particles that bind to targets. The isolates can be grown up overnight, and the displayed peptide sequence
15 responsible for binding can be deduced by DNA sequencing.

These libraries have approximately >10^5 different recombinants, and nucleotide sequencing of the inserts suggests that the expressed peptides are indeed random in amino acid sequence. These libraries are referred to herein
20 as TSAR libraries, where TSAR stands for Totally Synthetic Affinity Reagents. The preparation of the TSAR libraries are described further below.

5.3. SH3 Binding Clones And Their Characteristics

25   Accordingly, peptides have been isolated from an unconstrained random peptide library which exhibit a binding affinity for SH3 domains. Furthermore, the binding affinities exhibited by the disclosed peptides differ in their selectivities with certain peptides showing comparable
30 binding affinities for SH3 domains derived from different proteins, while others manifest greater affinities for specific SH3 domains.

The amino acid sequence of various peptides isolated by the present method are listed in Figure 5. As can be seen
35 from this list, c rtain groups of SH3 domain binding peptides are is lat d from three separat random peptide librari s, each bas d on a different type of random p ptide insert, all

- 22 -

displayed at the amin -terminus of the pIII prot in on the
surface of M13 viral particles. T n clon s w re isolated
from the R8C library, seven from the TSAR-12 library, and
seven from the TSAR-9 library. The sequences are presented
5 to highlight the particular amino acid residues believed to
bind directly to the SH3 domain, as well as to point out the
remaining amino acid resiudes of the random insert and the
viral flanking sequences and complementary site amino acid
residues common to each group of clones. The frequency with
10 which each particular clone is found in each library is also
indicated in Figure 5. Thus, clones T12.SRC3.1 and
T12.SRC3.2 are by far the most abundant clones found among
the three libraries.

Interestingly, all the binding peptides are found to
15 have the proline-rich amino acid residue motif, which is
apparently responsible for binding, the motif being located
predominantly at the C-terminal end of the insert, although
each clone also contains an insert at the N-terminal end.
The significance of this observation is not presently
20 understood, although this finding may indicate the possible
importance of the C-terminal viral flanking sequences in SH3
domain binding.

Indeed, a synthetic peptide bearing only the core
consensus sequence RPLPPLP (SEQ ID NO:9) was less effective
25 in binding to target SH3 domains than synthetic peptides that
also included additional amino acid residues flanking the
core sequences. Thus, 13-mers and 14-mers having the
sequences RSTPRPLPMLPTTR (SEQ ID NO:62), RSTPRPLPPLPTTR (SEQ
ID NO:67), GILAPPVPPRNTR (SEQ ID NO:63), GPHRRLPPTPATR (SEQ
30 ID NO:65), and VLKRPLPIPPVTR (SEQ ID NO:64) have been
prepared and shown to bind to SH3 domains, such as those of
Src and Yes, much more avidly than the 7-mer, RPLPPLP (SEQ ID
NO:9). The 13-mer ANPSPATRPLPTR (SEQ ID NO:66) has been
shown to have binding affinities comparable to the core
35 consensus sequenc . In each case, the 13-mers comprise a 7-
mer "core" sequence plus additional amino acid residues

- 23 -

SUBSTITUTE SHEET (RULE 26)

flanking same, some of which additional amino acid residues
are contributed by the viral flanking sequences.

Thus, in one embodim nt of the present invention, a 7-
mer core includes a consensus motif of the formula RXLPφφP
5 (SEQ ID NO:71), wherein R is arginine, L is leucine, P is
proline, X represents any amino acid except cysteine and φ
represents a hydrophobic amino acid residue. By "hydrophobic
amino acid residue," the applicants mean to include F, Y, W,
V, A, I, L, P or M, each letter representing the standard
10 one-letter designation for the corresponding amino acid
residue.

Furthermore, a preferred 9-mer peptide comprising two
additional amino acids on the C-terminal end of the core
sequence is envisioned having a consensus motif of the
15 formula RXLPφφPXψ. In this preferred 9-mer consensus motif,
the symbol ψ represents a hydrophilic amino acid residue,
except cysteine. By "hydrophilic amino acid residue," the
applicants mean to include K, R, H, D, E, N, Q, T, S or C,
and the other symbols are as defined above. For the purposes
20 of the present invention, a glycine residue (G) may be
considered either a hydrophobic or a hydrophilic amino acid
residue. The one-letter symbols B and Z, which stand for N
or D and Q or E, respectively, are considered hydrophilic
amino acid residues.

25   Particular 13-mer peptides of the present invention
include those listed, below. It is noted, however, that not
all the following 13-mer peptides correlate strictly to or
comply with the preferred 9-mer consensus motif, described
above. Those peptides that *do not* comply (indicated in
30 italics, with the non-complying amino acid residues
underscored) can, thus, be described as "resembling" those
that do comply (indicated in normal type) with the preferred
9-mer consensus motif: PGFRELPPLPPSR (SEQ ID NO:72),
AQSRPLPIPPETR (SEQ ID NO:73), VLKRPLPIPPVTR (SEQ ID NO:64),
35 *PPNSPLPPLPTHL* (SEQ ID NO:74), *TGRGPLPPLPNDS* (SEQ ID NO:75),
YSTRPVPPITRPS (SEQ ID NO:76), *SHKSRLPPLPTRP* (SEQ ID NO:77),
YRFRALPSPPSAS (SEQ ID NO:78), *GPHRRLPPTPATR* (SEQ ID NO:65),

- 24 -

LAQRQLPPTPGRD (SEQ ID NO:79), _ALQRRLPRTPPPA_ (SEQ ID NO:80),
PATRPLPTRPSRT (SEQ ID NO:81), _YSTRPLPSRPSRT_ (SEQ ID NO:82),
XPGRILLLPSEPR (SEQ ID NO:83), _SGGILAPPVPPRN_ (SEQ ID NO:84),
RSTRPLPILPRTT (SEQ ID NO:85), STPRPLPMLPTTR (SEQ ID NO:86),

5 STNRPLPMIPTTR (SEQ ID NO:87), _RSTRPLPSLPITT_ (SEQ ID NO:88),
STSRPLPSLPTTR (SEQ ID NO:89), RSTRSLPPLPPTT (SEQ ID NO:90),
RSTRQLPIPPTTT (SEQ ID NO:91), STPRPLPLIPTTP (SEQ ID NO:92),
RSTRPLPPTPLTT (SEQ ID NO:93), and _RSTRPQPPPPITT_ (SEQ ID
NO:94). Accordingly, other peptides not specifically

10 disclosed, which either comply with or "resemble" the
preferred 9-mer consensus motif, can be readily envisioned by
those of ordinary skill in the art and are considered to be
equivalent to those that are specifically disclosed above.
In particular, non-compliance at positions 1 (S, G, and I, in

15 place of R, are tolerated), 3 (V, A, and Q, in place of L,
are tolerated), 4 (L, in place of P, is tolerated), 5
(hydrophilic amino acid residues, S, R, and T, are tolerated
in place of a hydrophobic amino acid residue), 6 (hydrophilic
amino acid residues, R and T, are tolerated in place of a

20 hydrophobic amino acid residue), 7 (T, and S, in place of P,
are tolerated), and 9 (P and A are tolerated in place of a
hydrophilic amino acid residue) have been observed.

5.3.1.   Binding Specificities

25   It has been discovered that certain of the
binding peptides disclosed have a greater relative binding
affinity for one SH3 domain over another. Referring now to
Figure 8, the relative binding affinities of the various
peptides described above toward different SH3 domain targets

30 are graphically presented. As one can see, the relative
binding affinities of the respective peptides can differ by
orders of magnitude. Thus, as shown in Figure 8, the peptide
GPHRRLPPTPATR (SEQ ID NO:65), having the relevant sequence of
the phage clone identified as T12.SRC3.3, is specific to Src

35 family SH3 domains, including, but not limited to, Src, Yes,
Lck, Hck, Fgr, Fyn, and Lyn. This SH3 binding peptide has
little affinity for SH3 domains derived from PLCγ or Grb2.

- 25 -

SUBSTITUTE SHEET (RULE 26)

On th other hand, the peptide GILAPPVPPRNTR (SEQ ID NO:63),
corr sponding to the r levant s quence of the phage clone
T12.SRC3.1, which is one of the m st abundant binding clones
found by the present method, binds generically to a broad
5 range of SH3 domains, including Src, PLCγ, and Grb2.

On an intermediate level, the present invention has also
uncovered a peptide, VLKRPLPIPPVTR (SEQ ID NO:64),
corresponding to the relevant sequence of the phage clone
T12.SRC3.6, which is Src preferential; that is, this peptide
10 exhibits strong binding affinities for members of the Src
family, some binding affinities for Grb2 proteins, but littl
binding affinities for PLCγ domains. The peptide
ANPSPATRPLPTR (SEQ ID NO:66), corresponding to the relevant
sequence of the phage clone T12.SRC3.2, also exhibits Src
15 family specificity similar to GPHRRLPPTPATR (SEQ ID NO:65).
The peptides RSTPRPLPMLPTTR (R8C.YES3.5; SEQ ID NO:62) and
RSTPRPLPPLPTTR (representative consensus motif; SEQ ID NO:67)
are highly specific for SH3 domain of Src, Yes, and other
Src-related proteins.

20

5.4. Further Discussion of Binding Experiments

At the outset it is apparent that the binding
affinity of certain peptides to the SH3 domain of Src and
Src-related proteins is governed by more than just the
25 presence of the preferred core consensus sequences, RPLPPLP
(SEQ ID NO:9) or RPLPMLP (SEQ ID NO:95; i.e., RPLP(P/M)LP,
SEQ ID NO:96). Thus, while the synthetic peptides
RSTPRPLPMLPTTR (R8C.YES3.5; SEQ ID NO:62) and RSTPRPLPPLPTTR
(consensus; (SEQ ID NO:67) exhibit a strong specific binding
30 affinity for Src SH3, the other synthetic peptides tested
also exhibited an avid binding affinity to SH3 domains
relative to the 7-mer, RPLPPLP (SEQ ID NO:9). These other
peptides, GILAPPVPPRNTR (SEQ ID NO:63), VLKRPLPIPPVTR (SEQ ID
NO:64), GPHRRLPPTPATR (SEQ ID NO:65), and ANPSPATRPLPTR (SEQ
35 ID NO:66), sport core sequences and flanking sequenc s that
do not closely adhere to the preferred core consensus
sequences. Thus, these results suggest that binding affinity

- 26 -

to SH3 domains is gov rned to a large extent by the nature of
the amino acid residu s flanking the core 7-mer sequence.

        The binding characteristics of Src SH3-selected peptides
was determined using synthetic biotinylated peptides
5 corresponding to the sequences displayed by Src SH3-selected
phage.   These biotinylated peptides were assayed for direct
binding to immobilized Src SH3-GST.   Each of the five
library-derived peptides tested were found to bind to Src
SH3-GST and Yes SH3-GST over background (Figure 8).
10 Furthermore, a strong correlation was observed between the
similarity of a given peptide to the preferred core consensus
sequence RPLP(P/M)LP and the peptide's affinity for Src SH3-
GST.   The core sequence of the clone T12.SRC3.1
(GILAPPVPPRNTR; SEQ ID NO:63) appears to provide more generic
15 SH3 domain-binding characteristics.

        Experiments comparing the relative binding of various
phage clones to SH3 domains taken from a variety of proteins
demonstrated the preference of these clones for Src and Src-
related SH3 domains over SH3 domains taken from other
20 proteins.

        It was further found that while the 7-mer having the
consensus sequence RPLPPLP (SEQ ID NO:9) bound to Src SH3-GST
only weakly, peptides comprising the consensus sequence
flanked by residues encoded by one of the Src SH3-selected
25 clones (R8C.YES3.5), RSTP (SEQ ID NO:97) at the N-terminal
end and TTR at the C-terminal end, bound significantly better
than any of the peptides tested (Figure 7).   Thus, as stated
previously, sequences that flank the RPLP(P/M)LP (SEQ ID
NO:96) core appear to be important contributors to SH3
30 binding.   It is further surmised that a peptide having or
resembling the sequence RSTPAPPVPPRTTR (SEQ ID NO:98) should
exhibit strong but generic binding to a variety of SH3
domains.

        Similarly, it is observed that most of the Src SH3-
35 binding motifs are located near the carboxy-terminus of the
random peptides, adjacent to sequences which are fixed in
every clone (Figur 5).   The exceptional clones tend to

- 27 -

SUBSTITUTE SHEET (RULE 26)

poss ss sequenc s that resemble motifs that include fixed
flanking s quences. This clustering contrasts with previ us
results, in which binding motifs are distributed throughout
the random peptide. Kay, B.K., et al., in _Gene_ (1993)
5 128:59-65.

The binding of the library-derived Src SH3-binding
peptides was compared to that of peptides corresponding to
proline-rich regions of natural proteins. Peptides
corresponding to SH3-binding regions in human PI-3' Kinase
10 (KISPPTPKPRPPRPLPV; SEQ ID NO:69) and human SOS1.20
(GTVEPVPPPVPPRRRPESA; SEQ ID NO:68), as well as a proline-
rich region of the cytoskeletal protein vinculin
(LAPPKPPLPEGEV; SEQ ID NO:70), bound Src SH3 less well than
the library-derived peptides (Figure 7).

15   As mentioned above, the relative specificity of binding
was explored. Thus, the relative binding of Src SH3-select d
peptides to equal amounts of GST fusions to SH3 domains from
different proteins was determined (Figure 8). While all of
the library-derived peptides bound the Src and Yes SH3
20 domains almost equally well, none of the peptides (with the
exception of peptide T12.SRC3.1, the most divergent peptide
tested) bound the SH3 domains of Grb2, Crk, Abl or PLCγ1
appreciably. Thus, the library-derived peptides, in contrast
with a peptide derived from SOS1, exhibit SH3 binding that is
25 relatively specific for Src-family members.

Next, it was determined whether the binding to the Src
SH3 domain was qualitatively like the interactions of the SH3
domain and natural proteins found in cell lysates. Thus,
radiolabeled proteins were prepared from NIH 3T3 cell lysates
30 and chromatographed over Src SH3-GST immobilized on
glutathione linked Sepharose. SDS-PAGE shows that a number
of proteins can be affinity purified in this manner. The
synthesized peptides bind quite well to the Src SH3 domain,
as they can compete the binding of radiolabeled proteins from
35 cell lysates to immobiliz d Src-GST, with an IC_{50} of 1-10 mM
(Figure 9). In conclusion, the peptides can efficiently

- 28 -

SUBSTITUTE SHEET (RULE 26)

block the interaction of cellular proteins with Src SH3 *in vitro*.

Moreover, *Xenopus laevis* oocytes injected with mRNA encoding constitutively active Src undergo progesterone-
5 induced maturation at an accelerated rate relative to oocytes injected with water or c-Src mRNA. Unger, T.F. and Steele, R.E. in Mol. Cell.Biol. (1992) 12:5485-5498. To explore the ability of the library-derived Src SH3-binding peptides to exert a biochemical effect *in vivo*, the influence of the
10 peptides on the maturation of *Xenopus laevis* oocytes was examined. Hence, stage VI oocytes were injected with peptide, exposed to progesterone, and scored for germinal vesicle breakdown. Figure 10 shows that the rate of maturation was accelerated by approximately one hour when
15 oocytes were injected with the SH3-binding peptide consisting of RPLPPLP (SEQ ID NO:9) flanked by residues from clone T12.SRC3.6 (VLKRPLPIPPVTR; SEQ ID NO:64), but not with water or a peptide corresponding to a proline-rich segment of vinculin (LAPPKPPLPEGEV; SEQ ID NO:70) as controls. The
20 magnitude of this effect is roughly equivalent to that seen with injection of mRNA encoding constituitively active Src. See, e.g., Figure 3B in Unger, T.F. and Steele, R.E., *supra*. This result suggests that the library-derived Src SH3-binding peptide is effectively relieving an inhibitory effect of the
25 Src SH3 domain upon Src PTK activity. This model is consistent with a number of studies which have demonstrated an inhibitory effect of the Src SH3 domain upon Src kinase and transforming activity. See, e.g., Okada, M., et al., *supra*; Murphy, S.M., et al., *supra*; and Superti-Furga, G., et
30 al., *supra*.

5.5. Diagnostic And Therapeutic Agents
         Based On SH3 Binding Peptides and Additional
Methods   of Their Use

As already indicated above, the present invention
35 also seeks to provide diagnostic, prophylactic, and

- 29 -

therap utic agents based on the SH3 binding peptides
described herein.

In one embodiment, diagnostic agents are provided,
preferably in the form of kits, comprising an SH3 domain-
5 binding peptide and a detectable label conjugated to said
peptide directly, indirectly or by complexation, said peptide
comprising: (i) a core sequence motif of the formula RXLP$\phi\phi$P
(SEQ ID NO:71), wherein X represents any amino acid except
cysteine and ϕ represents a hydrophobic amino acid residue,
10 including F, Y, W, V, A, I, L, P, M or G, each letter
representing the standard one-letter designation for the
corresponding amino acid residue; and (ii) two or more
additional amino acid residues flanking said core sequence at
its C-terminal end, N-terminal end or both.
15   The diagnostic agents of the present invention can be
used to detect the presence of SH3 domains of a generic or
specific type in cells, tissues or organs either *in vitro* or
in vivo. For *in vivo* applications, the diagnostic agent is
preferably mixed with a pharmaceutically acceptable carrier
20 for administration, either enterally, parenterally or by some
other route dictated by the needs of the particular
application.

In a particular embodiment, for example, an assay based
on a fusion product is contemplated which comprises a Src SH3
25 domain-binding peptide of the invention and a substrate for
deregulated or "activated" Src. For instance, a muscle
biopsy, taken from a subject suspected of being infected by
the Rous sarcoma virus, can be treated with an effective
amount of the fusion product. By subsequent analysis of the
30 degree of conversion of the substrate, one can potentially
detect infection by the Rous sarcoma virus in the subject,
particularly mammals, especially chickens. The presence of
the retrovirus, which causes the expression of deregulated or
"activated" Src, may thus be indicated by unusually high
35 levels of Src as revealed by large amounts of the converted
substrate. Se , for xample, Paxton, W.G. et al., in
Biochem. Biophys. Res. Commun. (1994) 200(1):260-267

- 30 -

SUBSTITUTE SHEET (RULE 26)

(d tection of phosphorylated tyrosine and serine residu s of angiotensin II AT1 r ceptor, a ubstrate of Src family tyrosine kinases); another suitable substrate may be th protein p68 (Fumagalli, S. et al., in <u>Nature</u> (1994)

5 368(6474):871-874; Taylor, S.J. and Shalloway, D., in <u>Ibid.</u> at 867-871.

Alternatively, the enzyme can be isolated by selective binding to a form of the SH3 domain-binding peptides of the present invention (e.g., biotin-peptide conjugate). After

10 isolation of the protein-peptide conjugate complex (e.g., on a column comprising streptavidin), the activity of the enzyme can then be assayed by conventional methods to determine its level of protein kinase activity which can be taken as an indication of the presence of the deregulated or "activated"

15 form of the enzyme. An assay for Src kinase has been described by Klinz and Maness, in <u>Neuroprotocols</u> (a companion to <u>Neuroscience</u>)(1992) 1(3):224-231.

Moreover, the diagnostic agents of the invention can also serve as imaging agents of cells, tissues or organs,

20 especially those that contain proteins with an SH3 domain. For example, neural cells (e.g., neurons, other areas of the brain), osteoclasts, osteoblasts, platelets, immune cells, and other dividing cells are known to express or contain proteins with SH3 domains. Thus, an image can be taken of

25 portions of the body to serve as a baseline for subsequent images to detect physiologic or biochemical changes in the subject's body. For instance, changes in the condition of cellular levels of Src or a transformation of the cellular Src to an "activated" form may be detected using the

30 diagnostic or imaging agents of the present invention.

Accordingly, it has been demonstrated that an SH3-binding peptide tagged with a fluorescence emitter can provide an image of the cytoskeleton. The images are presented in Figure 11. As can be seen from Figure 11,

35 pan ls A, B, and C show the fluorescence image that is obtain d on treating NIH 3T3 fibroblasts with SH3 domain-binding peptides modified to include a fluorescent tag. In

- 31 -

sharp contrast, panel D shows only a dark image that is produced when the cells are treated with a proline-rich segment of vinculin as a control.

In another embodiment, an SH3 domain-binding peptide-
5 horseradish immunoperoxidase complex or related immunohistochemical agent could be used to detect and quantitate specific receptor molecules in tissues, serum or body fluids. In particular, the present invention provides useful diagnostic reagents for use in immunoassays, Southern
10 or Northern hybridization, and *in situ* assays. Accordingly, the diagnostic agents described herein may be suitable for use *in vitro* or *in vivo*.

In addition, the diagnostic or imaging agent of the present invention is not limited by the nature of the
15 detectable label. Hence, the diagnostic agent may contain one or more such labels including, but not limited to, radioisotope, fluorescent tags, paramagnetic substances, heavy metals, or other image-enhancing agents. Those of ordinary skill in the art would be familiar with the range of
20 label and methods to incorporate or conjugate them into the SH3 domain-binding peptide to form diagnostic agents.

In yet a further embodiment, pharmaceutical compositions are provided comprising an SH3 domain-binding peptide and a pharmaceutically acceptable carrier. In a specific
25 embodiment of the invention, the pharmaceutical composition is useful for the modulation of the activity of SH3 domain-containing proteins. By "modulation" is meant either inhibition or enhancement of the activity of the protein target. Accordingly, a pharmaceutical composition is
30 disclosed comprising an SH3 domain-binding peptide and a pharmaceutically acceptable carrier, said peptide comprising: (i) a 9-mer sequence motif of the formula RXLP$\phi\phi$PXψ (SEQ ID NO:10), wherein X represents any amino acid except cysteine, ϕ represents a hydrophobic amino acid residue, and wherein ψ
35 is a hydrophilic amino acid residue except cysteine, each lett r representing the standard one-lett r designation for the corresponding amino acid residu ; and, optionally, (ii)

- 32 -

SUBSTITUTE SHEET (RULE 26)

additional amino acid residues flanking the 9-mer s quence at its C-terminal end, N-terminal end or both, up to a total of 45 amino acid residues, including said 9-mer sequence. Preferably, the peptide comprises at least one, more

5 preferably at least two, and most preferably at least three additional amino acids flanking the 9-mer sequence.

As stated above, the therapeutic or diagnostic agents of the invention may also contain appropriate pharmaceutically acceptable carriers, diluents and adjuvants. Such

10 pharmaceutical carriers can be sterile liquids, such as water and oils including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered

15 intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel,

20 magnesium carbonate, magnesium stearate, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. These compositions can take the form of solutions, suspensions, tablets, pills, capsules, powders, sustained-

25 release formulations and the like. Suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin.

Such compositions will contain an effective therapeutic amount of the active compound together with a suitable amount

30 of carrier so as to provide the form for proper administration to the subject. While intravenous injection is a very effective form of administration, other modes can be employed, including but not limited to intramuscular, intraperitoneal, and subcutaneous injection, and oral, nasal,

35 enteral, and parenteral administration.

The therapeutic agents and diagnostic agents of th instant inv ntion are used for the tr atment and/or diagnosis

- 33 -

of animals, and more preferably, mammals including humans, as
well as dogs, cats, hors s, cows, pigs, guinea pigs, mic and
rats. Accordingly, other methods contemplated in the pres nt
invention, include, but are not limited to, a method of
5 modulating, i.e., inhibiting or enhancing, bone resorption in
a mammal (see, e.g., Hall, T.J., in Biochem. Biophys. Res.
Commun. (1994) 199(3):1237-44), a method of disrupting
protein tyrosine kinase-mediated signal transduction pathways
or a method of regulating the processing, trafficking or
10 translation of RNA in a cell by introducing or administering
an effective amount of an SH3 domain-binding peptide of the
present invention (see, e.g., Taylor, S.J. and Shalloway, D.,
supra).

The diagnostic or therapeutic agents of the present
15 invention can be modified by attachment to soluble
macromolecules such as proteins, polysaccharides, or
synthetic polymers. For example, the peptide could be
coupled to styrene-maleic acid copolymers (see, e.g.,
Matsumura and Maeda, Cancer Res. (1986) 46:6387),
20 methacrylamide copolymers (Kopececk and Duncan, J. Controlled
Release (1987) 6:315), or polyethylene glycol (PEG) (e.g.,
Hershfield and Buckley, N. Engl. J. Med. (1987) 316:589; Ho
et al., Drug Metab. Dispos. (1986) 14:349; Chua et al., Ann.
Intern. Med. (1988) 109:114).   The agents, if desired,
25 are further targeted by attachment to an antibody, especially
a monoclonal antibody. Such antibodies include but are not
limited to chimeric, single chain, Fab fragments, and Fab
expression libraries. In one embodiment the agent is coupl d
to the macromolecule via a degradable linkage so that it will
30 be released in vivo in its active form.

In another embodiment, the therapeutic or diagnostic
agent may be delivered in a vesicle, in particular a
liposome. See, Langer, Science (1990) 249:1527-1533; Treat
et al., in Liposomes in the Therapy of Infectious Disease
35 and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York
(1989) pp. 353-365; Lopez-Berestein, ibid., pp. 317-327.

- 34 -

SUBSTITUTE SHEET (RULE 26)

In yet another embodiment, the therapeutic or *in vivo* diagnostic agent can b delivered in a controlled release system. In one embodiment, a pump may be used (se Langer, *supra*; Sefton, CRC Crit. Ref. Biomed. Eng. (1987) 14:201;

5 Buchwald et al., Surgery (1980) 88:507; Saudek et al., N. Engl. J. Med. (1989) 321:574). In another embodiment, polymeric materials may be used (see Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida, 1974; Controlled Drug Bioavailability, Drug

10 Product Design and Performance, Smolen and Ball (eds.) Wiley, New York 1984; Raner and Peppas, J. Macromol. Sci. Rev. Macromol. Chem. (1983) 23:61; see, also, Levy et al., Science (1985) 228:190; During et al., Ann. Neurol. (1989) 25:351; Howard et al., J. Neurosurg. (1989) 71:105). In a preferred

15 embodiment, a controlled release system may be placed next to the therapeutic target, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in Medical Applications of Controlled Release, *supra*, (1984) 2:115-138). It will be recognized by one of ordinary skill in the art that a

20 particular advantage of the invention is that a peptide will not be subject to the problems of denaturation and aggregation associated with proteins held in the warm, most environment of a body in a controlled release system.

Other controlled release systems are discussed in the

25 review by Langer, in Science (1990) 249:1527-1533.

30

35

- 35 -

SUBSTITUTE SHEET (RULE 26)

6.   **EXAMPLES**

    6.1. **Preparation of th TSAR-9 Library**

        6.1.1.   **Synth sis and Assembly of Oligonucleotides**

                Figure 1 shows the formula of the
oligonucleotides and the assembly scheme used in construction
of the TSAR-9 library. The oligonucleotides were synthesiz d
with an applied Biosystems 380a synthesizer (Foster City,
CA), and the full-length oligonucleotides were purified by
HPLC.

                Five micrograms of each of the pair of oligonucleotides
were mixed together in buffer (10 mM Tris-HCl, pH 8.3, 15 mM
KCl, 0.001% gelatin, 1.5 mM magnesium chloride), with 0.1 %
Triton X-100, 2 mM dNTP's, and 20 units of _Taq_ DNA
polymerase. The assembly reaction mixtures were incubated at
72 °C for 30 seconds and then 30 °C for 30 seconds; this
cycle was repeated 60 times. It should be noted that the
assembly reaction is not PCR, since a denaturation step was
not used. Fill-in reactions were carried out in a thermal
cycling, device (Ericomp, LaJolla, CA) with the following
protocol: 30 seconds at 72 °C, 30 seconds at 30 °C, repeated
for 60 cycles. The lower temperature allows for annealing of
the six base complementary region between the two sets of the
oligonucleotide pairs. The reaction products were
phenol/chloroform extracted and ethanol precipitated.
Greater than 90% of the nucleotides were found to have been
converted to double stranded synthetic oligonucleotides.

                After resuspension in 300 μL of buffer containing 10 mM
Tris-HCI, pH 7.5, 1 mM EDTA (TE buffer), the ends of the
oligonucleotide fragments were cleaved with _Xba_ I and _Xho_ I
(New England BioLabs, Beverly, MA) according to the
supplier's recommendations. The fragments were purified by
4% agarose gel electrophoresis. The band of correct size was
removed and electroeluted, concentrated by ethanol
precipitation and resuspended in 100 μL TE buffer.
Approximately 5% of the assembled oligonucleotides can be
expected to have internal _Xho_ I or _Xba_ I sites; however, only

- 36 -

SUBSTITUTE SHEET (RULE 26)

the full-length molecules were used in the ligation step of
the assembly scheme. The concentration of the synthetic
oligonucleotide fragments was estimat d by comparing the
intensity on an ethidium bromide stained gel run along with
5 appropriate quantitated markers. All DNA manipulations not
described in detail were performed according to Maniatis,
supra.

To demonstrate that the assembled enzyme digested
oligonucleotides could be ligated, the synthesized DNA
10 fragments were examined for their ability to self-ligate.
The digested fragments were incubated overnight at 18 °C in
ligation buffer with T4 DNA ligase. When the ligation
products were examined by agarose gel electrophoresis, a
concatamer of bands was visible upon ethidium bromide
15 staining. As many as five different unit length concatamer
bands (i.e., dimer, trimer, tetramer, pentamer, hexamer) were
evident, suggesting that the synthesized DNA fragments were
efficient substrates for ligation.

20   6.1.2.   **Construction of Vectors**
The construction of the M13 derived phage
vectors useful for expressing a TSAR library has been
recently described (Fowlkes, D. et al. BioTech. (1992)
13:422-427). To express the TSAR-9 library, an M13 derived
25 vector, m663, was constructed as described in Fowlkes. The
m663 vector contains the pIII gene having a *c-myc*-epitope,
i.e., as a stuffer fragment, introduced at the mature N-
terminal end, flanked by *Xho* I and *Xba* I restriction sites
(see also, Figure I of Fowlkes).
30
6.1.3.   **Expression of the TSAR-9 Library**
The synthesized oligonucleotides were then
ligated to *Xho* I and *Xba* I double-digested m663 RF DNA
containing, the pIII gene (Fowlkes) by incubation with ligase
35 overnight at 12 °C. More particularly, 50 ng of vector DNA
and 5 ng of the digest d synthesized DNA and was mixed
togeth r in 50 µL ligation buffer (50 mM Tris, pH 8.0, 10 mM

- 37 -

SUBSTITUTE SHEET (RULE 26)

MgCl$_2$, 20 mM DTT, 0.1 mM ATP) with T4 DNA ligase. After
overnight ligation at 12°C, the DNA was concentrated by
thanol precipitation and wash d with 70% thanol. The
ligated DNA was then introduced into *E. coli* (DH5αF'; GIBCO
5 BRL, Gaithersburg, MD) by electroporation.

        A small aliquot of the electroporated cells was plated
and the number of plaques counted to determine that 10^8
recombinants were generated. The library of *E. coli* cells
containing recombinant vectors was plated at a high density
10 (~400,000 per 150 mM petri plate) for a single amplification
of the recombinant phage. After 8 hr, the recombinant
bacteriophage were recovered by washing each plate for 18 hr
with SMG buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM
MgCl$_2$, 0.05% gelatin) and after the addition of glycerol to
15 50% were frozen at -80 °C. The TSAR-9 library thus formed
had a working titer of ~2 x 10^{11} pfu/ml.

6.2. Preparation of TSAR-12 Library

        Figure 2 shows the formula for the synthetic
20 oligonucleotides and the assembly scheme used in the
construction of the TSAR-12 library. As shown in Figure 2,
the TSAR-12 library was prepared substantially the same as
the TSAR-9 library described in Section 6.1 above with the
following exceptions: (1) each of the variant non-predicted
25 cligonucleotide sequences, i.e., NNB, was 30 nucleotides in
length, rather than 54 nucleotides; (2) the restriction sites
included at the 5' termini of the variant, non-predicted
sequences were *Sal* I and *Spe* 1, rather than *Xho* I and *Xba* I;
and (3) the invariant sequence at the 3' termini to aid
30 annealing of the two strands was GCGGTG and CGCCAC rather
than CCAGGT and GGTCCA (5' to 3').

        After synthesis including numerous rounds of annealing
and chain extension in the presence of dNTP's and Taq DNA
polymerase, and purification as described above in Section
35 6.1.1, the synth tic d uble strand d, oligonucleotide
fragments wer digested with *Sal* I and *Spe* I restriction
enzymes and ligated with T4 DNA ligase to the nucleotide

                            - 38 -

sequence encoding the M13 pIII gene contained in the m663
vector to yield a library of TSAR-expression vectors as
described in Sections 6.1.2 and 6.1.3. The ligated DNA was
then introduced into *E. coli* (DH5αF'; GIBCO BRL,
5 Gaithersburg, MD by electroporation. The library of *E. coli*
cells were plated at high density (~400,000 per 150 mm petri
plate) for amplification of the recombinant phage. After
about 8 hr, the recombinant bacteriophage were recovered by
washing, for 18 hr with SMG buffer and after the addition of
10 glycerol to 50% were frozen at -80 °C.

The TSAR-12 library thus formed had a working titer of
~2 x 10^{11} pfu/mL.

6.3. Characterization of the TSAR-9 and -12 Libraries

15   The inserted synthetic oligonucleotides for each of
the TSAR libraries, described in Sections 6.1 and 6.2 above,
had a potential coding complexity of 20^{36} (~10^{47}) and 20^{20},
respectively, and since ~10^{14} molecules were used in each
transformation experiment, each member of these TSAR
20 libraries should be unique. After plate amplification the
library solution or stock has 10^4 copies of each member/mL.

It was observed that very few (<10%) of the inserted
oligonucleotide sequences characterized so far in both of the
libraries have exhibited deletions or insertions. This is
25 likely a reflection of the accuracy assembling the
oligonucleotides under the conditions used and the fact that
certain types of mutations (i.e., frame-shifts) would not be
tolerated as pIII an essential protein for phage propagation.

In order to determine whether any coding bias existed in
30 the variant non-predicted peptides expressed by these
libraries, perhaps due to biases imposed *in vitro* during
synthesis of the oligonucleotides or *in vivo* during
expression by the reproducing phage, inserts were sequenced
as set forth below.

35

6.3.1.   Charact rization of TSAR-9 Library

- 39 -

SUBSTITUTE SHEET (RULE 26)

Insert d synthetic oligonucleotide fragments
of 23 randomly chosen isolates were examined from the TSAR-9
library. Individual plaques were used to inoculate I ml of
2XYT broth containing *E. coli* (DH5αF') cells and the cultures
5 were allowed to grow overnight at 37°C with aeration. DNA
was isolated from the culture supernatants according to
Maniatis, *supra*. Twenty-three individual isolates were
sequenced according to the method of Sanger (<u>Proc. Natl.</u>
<u>Acad. Sci. USA</u> (1979) 74:5463-5467) using as a primer the
10 oligonucleotide 5'-AGCGTAACGATCTCCCG (SEQ ID NO. 99), which
is 89 nucleotides downstream of the pIII gene cloning site of
the m663 vector used to express the TSARS.

Nucleotide sequences and their encoded amino acid
sequences were analyzed with the MacVector computer program
15 (IBI, New Haven, CT). The Microsoft EXCEL program was used
to evaluate amino acid frequencies. Such analyses showed
that the nucleotide codons coding for and hence most amino
acids, occurred at the expected frequency in the TSAR-9
library of expressed proteins. The notable exceptions were
20 glutamine and tryptophan, which were over- and under-
represented, respectively.

It is of interest to note the paucity of TAG stop codons
in the inserts, i.e., only 2 of ~200 isolates characterized
contained a TAG stop codon. About half $[1-(47/48)^{36}]$ of the
25 phage inserts were expected to have at least one TAG codon in
view of the assembly scheme used. However, most of the TAG-
bearing phage appear to have been lost from the library, even
though the bacterial host was *supE*. This may be a
consequence of suppression being less than 100% effective.
30   The amino acids encoded by the inserted double strand d
synthesized oligonucleotide sequences, excluding the fixed
PG-encoding centers, were concatenated into a single sequenc
and the usage frequency determined for each amino acid using
the Microsoft EXCEL program. These frequencies were compared
35 to that expected from the assembly scheme of the
oligonucl otides, and the divergence from expect d values
repres nt d by the size of the bars above and below the

- 40 -

SUBSTITUTE SHEET (RULE 26)

baseline. Chi square analysis was used to d termine the
significance of the deviati ns. The majority of amino acids
were found to occur at the expected fr quency, with th
notable exceptions that glutamine and tryptophan were
5 somewhat over- and under-represented, respectively. Thus,
except for the invariant Pro-Gly, any position could have any
amino acid; hence, the sequences are unpredicted or random.

6.3.2.   Characterization of TSAR-12 Library

10   Approximately 10 randomly chosen inserted
oligonucleotides from the TSAR-12 library were examined by
DNA sequencing as described above in Section 6.3.1. The
isolates were chosen at random from the TSAR-12 library and
prepared for sequencing, as were the TSAR-9 isolates.
15 Analysis showed that except for the invariant Gly any
position could have any amino acid; hence, the sequences are
unpredicted or random.

6.4. Preparation of R8C Library

20   Referring now to Figure 3, two oligonucleotides
were synthesized on an Applied Biosystems Model 380a machine
with the sequence 5'-
TGACGTCTCGAGTTGTNNKNNKNNKNNKNNKNNKNNKNNKTGTGGATCTAGAAGGATC-3'
(SEQ ID NO:31) and 5'-GATCCTTCTAGATCC-3' (SEQ ID NO:32),
25 where N is an equimolar ratio of deoxynucleotides A, C, G,
and T, and K is an equimolar ratio of G and T. Fifty pmol of
each oligonucleotide was incubated at 42 °C for 5 min, then
37 °C for 15 min, in 50 μL of Sequenase™ buffer (U.S.
Biochemicals, Cleveland, OH) with 0.1 μg/μL acetylated BSA,
30 and 10 mM DTT. After annealing, 10 units of Sequenase™ (U.S.
Biochemicals) and 0.2 mM of each dNTP were added and
incubated at 37 °C for 15 min. The sample was then heated at
65 °C for 2 hr, digested with 100 units of both Xho I and Xba
I (New England BioLabs, Beverly, MA), phenol extracted,
35 ethanol precipitated, and resolved on a 15% non-denaturing
polyacrylamide gel. The assembled, digested fragment was gel
purified prior to ligation. The vector, m663 (Fowlkes, D. t

- 41 -

SUBSTITUTE SHEET (RULE 26)

al. <u>Biotech</u>. (1992) 13:422-427), was prepared by dig stion
with *Xho* I and *Xba* I, calf alkaline phosphatase (Bo hring r
Mannheim, Indianapolis, IN) treatm nt, phenol extracted, and
purified by agarose gel electrophoresis. To ligate, 20 μg
5 vector was combined with 0.2 μg insert in 3 mL with T4 DNA
ligase (Boehringer Mannheim), according to the manufacturer.
After removal of the protein and buffer by phenol extraction
and ethanol precipitation, the ligated DNA was electroporated
into XL1-Blue *E. coli* (Stratagene, San Diego, CA) and plated
10 for eight hours at 37 °C. To recover the recombinant phage,
the top agar was collected with a spatula, mixed with an
equal volume of 100 mM NaCl, 10 mM $MgCl_2$, and 50 mM Tris-HCI
(pH7.5), and disrupted by two passes through an 18-gauge
syringe needle. The bacterial cells were removed by
15 centrifugation, and phage particles were collected by
polyethylene glycol precipitation and stored at -70 °C in 25%
glycerol. The library had 10^8 total recombinants and a
working titer of 6 x 10^{13} pfu/mL.

Members of the library were checked for inserts by th
20 polymerase chain reaction (Saiki, et al. <u>Science</u> (1988)
239:487-491). Individual plaques on a petri plate were
touched with a sterile toothpick and the tip was stirred into
2xYT with F^+ *E. coli* bacteria and incubated overnight at 37 °C
with aeration. Five microliters of the phage supernatant
25 were then transferred to new tubes containing buffer (67 mM
Tris-HCl, pH 8.8/10 mM β- mercaptoethanol/16.6 mM ammonium
sulfate/6.7 mM EDTA/50 μg bovine serum albumin per mL), 0.1
mM deoxynucleotide triphosphates, and 1.25 units of *Taq* DNA
polymerase (Boehringer Mannheim, Indianapolis, IN) with 100
30 pmoles of oligonucleotide primers. The primers flanked the
cloning site in gene III of m663 (5'-TTCACCTCGAAAGCAAGCTG-3'
(SEQ ID NO:100) and
5'-CCTCATAGTTAGCGTAACG-3' (SEQ ID NO:101)). The assembly
reactions were incubated at 94 °C for 1 min, 56 °C for 2 min,
35 and 72 °C for 3 min; this cycle was repeated 24 times. Th
reaction products wer then resolved by electrophoresis on a
NuSieve 2.0% agaros g l (FMC, Rockland, ME). Gels reveal d

- 42 -

SUBSTITUTE SHEET (RULE 26)

that for 20 plaques tested, all were recombinant and had
single inserts of the xpected size.

Based on the sample size of the library, it was
anticipated that 100% of the recombinants had single inserts.
5 However, all of the SH3-binding phage isolated from the R8C
library had double-inserts. Such phage are presumed rare
(i.e., <5%) within the library, yet because the SH3-binding
peptide appears to need to be linear they were selected for
by our screening methods. Most likely they were formed
10 during the generation of the library; one scenario is that
the inserts ligated together to form head-to-head dimers and
that they were subsequently cloned into m663 DNA by ligation
with the vector's *Xho I* sticky end and by illegitimate
ligation with the vector's *Xba I* site (see, Figure 4).

15

6.5. Preparation Of Target-Coated Microtiter Wells
6.5.1. Preparation Of GST-SH3 Fusion Proteins

The preparation of Src-GST fusion protein
was first described by Smith and Johnson, in <u>Gene</u> (1988)
20 67:31, the disclosure of which is incorporated by reference
herein. Briefly, pGEX-derived (Pharmacia, Piscataway, NJ)
constructs expressing GST fusion proteins containing the SH3
domains of Src, Grb2, Crk, Abl, or PLCγ were obtained from
Dr. Channing Der (University of North Carolina at Chapel
25 Hill); a construct expressing the SH3 domain of Yes was
obtained from Dr. Marius Sudol (Rockefeller University). The
use of the pGEX bacterial expression vector for the
production of GST-SH3 fusion proteins is well-known to those
in the art. See, e.g., Cicchetti, P. et al., in <u>Science</u>
30 (1992) 257:803-806. Briefly, the coding region for a
particular SH3 domain can be fused in-frame at the Bam HI
site of pGEX-2T. Thus, fusion proteins were prepared as per
the manufacturer's instructions, and quantified by Coomassie
Blue staining of SDS-polyacrylamide gels. Microtiter wells
35 were coated with 5-20 μg GST-SH3 fusion protein in 100 mM
NaHCO$_3$, pH 8.5, blocked with 100 mM NaHCO$_3$ (pH 8.5) 1% BSA,
and washed. All washes consisted of five applications of

- 43 -

SUBSTITUTE SHEET (RULE 26)

1XPBS, 0.1% Tween 20, 0.1% BSA (Buffer A). Where
appropriate, the amount of protein bound to each well was
quantified with an anti-GST antibody-based ELISA (Pharmacia,
Piscataway, NJ), and with a GST-binding phage, isolated
5 during the course of this work.

     6.5.2.   Coating of Microtiter Wells

       Bacterially expressed Src SH3 glutathione-S-
transferase (Src-GST) fusion protein was purified from
bacterial lysates using glutathione agarose 4B (Pharmacia),
10 according to the manufacturer's instructions. Bound Src-GST
fusion protein was eluted from the glutathione agarose with
10 mM glutathione in PBS. Microtiter wells were then coated
with Src-GST fusion protein (1-10 μg/well, in 50 mM NaHCO$_3$, pH
8.5) overnight at 4 °C. To block non-specific binding of
15 phage, 100 μL 1% BSA in 100 mM NaHCO$_3$, pH 8.5, was added to
each well and allowed to incubate at room temperature for 1
hour. The wells were then washed five times with 200 μL PBS,
0.1% Tween 20, 0.1% BSA (Buffer A).

20   **6.6. Biopanning And Subsequent Characterization Of
      Phage-Displayed Random Peptide Libraries
      With Src-GST Fusion Protein As Target Molecule**

     6.6.1.   Isolation of Src SH3-Binding Phage

       Library screens were performed as
previously described. Kay, B.K., et al., in *Gene* (1993)
25 128:59-65. Briefly, 1 X 10^{11} pfu TSAR 9, TSAR 12, or R8C
phage in Buffer A were incubated in a Src SH3-GST-coated well
for 2 hours. The wells were washed, and bound phage were
eluted with 100 μL 50 mM glycine·HCl (pH 2.2), transferred to
a new well, and neutralized with 100 mL 200 mM NaHPO$_4$ (pH
30 7.0). Recovered phage were used to infect 1 x 10^9 DH5αF' *E.
coli* cells in 20 mL 2xYT; the infected cells were grown
overnight, resulting in a 1000- to 10,000-fold amplification
of phage titer. Amplified phage were panned twice more, as
above, excepting the amplification step. Binding phage
35 recovered after the third round of panning were plat d at a
low density on a lawn of DH5αF' *E. coli* cells to yield

- 44 -

isolated plaques for clonal analysis. Isolated plaques were used to produce small cultures from which phage stocks and DNA were recovered for phage binding experiments and dideoxy sequencing (Sanger, F., et al., in <u>Proc. Natl. Acad. Sci. USA</u>
5 (1977) 74:5463-5467), respectively. Clones were confirmed as binding the SH3 domain by applying equal titers of phage to wells containing Src SH3-GST or GST alone, and titering the number of eluted particles from each well, or detecting bound phage with an anti-phage antibody-based ELISA (Pharmacia).

10   Indeed, the ability of isolated phage clones to bind to several SH3 domains derived from a variety of different proteins can be investigated by the manner described above. GST-SH3 fusion proteins containing SH3 domains from a variety of different proteins are bound to microliter wells. An
15 aliquot of the aforementioned phage stocks (50 μL) is introduced into wells containing the different GST-SH3 fusion proteins. After room temperature incubation for 1-2 hours, the liquid contents of the microtiter plates are removed, and the wells are washed 5 times with 200 μL Buffer A. Bound
20 phage are eluted with 100 μL 50 mM glycine (pH 2.2), transferred to a new well, and neutralized with 100 μL 200 mM NaHPO$_4$ (pH 7.0). The phage are diluted 10^{-3}- to 10^{-6}-fold, and aliquots are plated onto lawns of DH5αF' *E. coli* cells to establish the number of plaque forming units in the output
25 sample. From these experiments, the relative specificity of different Src SH3 binding clones for SH3 domains derived from other proteins is determined.

6.6.2.   Phage ELISA and Nucleotide Sequencing

30   To evaluate the binding of isolates to various targets proteins, enzyme-linked-immuno-assays (ELISA) were also performed. Bacterial cultures were infected with phage isolates and cultured overnight in 2XYT at 37 °C. The cells were spun down and 25 mL of supernatant was added to
35 microtiter plate wells coated with 50 mL of protein (1 mg/mL in 100 mM NaHCO$_3$, pH 8.4; overnight at 4 °C or for a few hours at room temperature) and blocked (1 mg/mL BSA in 100 mM

- 45 -

NaHCO₃, pH 8.4; for about one hour). The phage are incubated
in the well with 25 mL of PBS-0.1% Tween 20 at RT for 2 hr.
The wells are then washed multipl times over 30 minut s. To
each well is added 50 μL of polyclonal anti-phage antibody
5   conjugated to horseradish peroxidase. The antibody is
diluted 1:3000 in BSA-PBS-Tween 20; it was obtained from
Pharmacia (Piscataway, NJ; catalog number 27-9402-01). Aft r
30 minutes, the wells are washed again with BSA-PBS-Tween 20
for ⁻20 minutes. Finally, 100 μL of ABTS reagent (Pharmacia,
10 with H₂O₂) are added to each well for the development of
color. Plates are read with a plate reader (Molecular
Devices, Menlo Park, CA) at 405 nm wavelength.

        The nucleotide sequence of the relevant segments of the
Src SH3 binding clones (or phage clones that bind to SH3
15 domains of other proteins) were sequenced using standard
methods. Sanger, F., et al., in _Proc. Natl. Acad. Sci. USA_
(1977) 74:5463-5467. The oligo primer 5'-AGCGTAACGATCTAAA-3'
(SEQ ID NO:102) was used, which is 89 nucleotides downstream
of the gene _III_ cloning site of M13 m666. The nucleotide
20 sequences were analyzed with the MacVector computer program
(IBI, New Haven, CT, USA). From this nucleotide sequence
information the primary sequence of each Src SH3 binding
peptide was deduced. The corresponding synthetic peptides
were then prepared by techniques well known in the art with
25 or without flanking sequences. Indeed, these synthetic
peptides have been shown to bind to SH3 domain targets, with
those possessing the phage flanking amino acid residues
exhibiting greater binding affinity.

30   **6.7 _In Vitro_ Peptide Binding Assays**
        Peptides were obtained from Research Genetics
(Birmingham, AL), Chiron Mimotopes (Victoria, Australia), or
synthesized by conventional techniques by Dr. J. Mark Carter
of Cytogen Corporation (Princeton, NJ). Peptide purity was
35 assessed by HPLC and/or mass spectrometry. Biotinylat d
peptides were synthesized with either a KSGSG (SEQ ID NO:103)
or a GSGS (SEQ ID NO:104) peptide linker (a spacer) betw en

- 46 -

the biotin and the N-terminus of the peptide. Binding
experiments wer performed as above, excepting the use of 10
μM peptide instead of phage. Bound biotinylated p ptide was
detected with streptavidin conjugated to alkaline phosphatase
5 (Sigma Chemical Co., St. Louis, MO).   After one hour
incubation period at room temperature, the wells were washed,
and a solution of 3 mM p-nitrophenyl-phosphate (US
Biochemicals, Cleveland, OH) in 50 mM $NaCO_3$ (pH 9.8), and 50
mM $MgCl_2$ was added and color allowed to develop. Signals were
10 read with an ELISA plate reader (Molecular Devices, Menlo
Park, CA) at 405 nm wavelength. Binding experiments were
performed in triplicate. The results are presented in
Figures 7 and 8.

15   **6.8. Peptide Competition of GST-SH3 Affinity
             Precipitations of Cell Lysates**
         Labeled proteins are prepared by incubating a
culture of HeLa cells overnight with ≥100 μCi/mL ^{35}S-
methionine. The cells are then washed and lysed with mild
20 detergent. This mixture of radioactive proteins is incubated
with Src-GST fusion protein that has been immobilized on
glutathione-linked Sepharose beads (Pharmacia, Piscataway,
NJ). After several hours of tumbling, the beads are pelleted
gently by low-speed centrifugation, and the supernatant is
25 discarded. The beads are then resuspended into a slurry in
PBS-0.1% Tween 20, pelleted, and washed several additional
times. Finally, a 2% SDS solution is added to the sample,
which is then boiled at 100 °C for 3 minutes. Afterward, the
sample is centrifuged, and the supernatant loaded on a 10%
30 polyacrylamide SDS gel for electrophoresis. After the
proteins have been resolved, the gel is fixed, dried down,
and exposed to X-ray film for autoradiography or phosphor
plates for scanning by a Molecular Dynamics PhosphorImager.
         The ability of Src SH3 to bind certain ^{35}S-labeled
35 proteins is examined for competability with exogenous
peptides. Synthetic peptides corresponding to phage-
displayed ins rts and motifs are added at the time that the

- 47 -

SUBSTITUTE SHEET (RULE 26)

lysat is incubated with the Src-GST fusion protein
immobilized on glutathione-linked sepharose beads. The SH3
binding peptides block binding of all or some of the labeled
proteins while negative control peptides (unrelated peptide
5 sequences) do not. The amount of competition is quantified
and correlated with the amount of added SH3-domain binding
peptides.

      Alternatively, NIH 3T3 cells were grown in Dulbecco's
Modified Eagle Medium (DME) + 10% fetal calf serum (FCS) + 80
10 μCi/mL Tran^{35}Slabel (ICN), washed with PBS, lysed in RIPA
buffer, and pelleted. Supernatant from 1.5×10^6 cells was
precleared with 100 μg glutathione-agarose-immobilized GST.
The supernatant was then incubated with 10 μg glutathione-
agarose-immobilized GST-SH3 fusion protein with or without
15 added test peptide in a final volume of 250 μL. Pelleted
beads were washed with 1 mL each of RIPA, RIPA + 1%
deoxycholate + 0.1% SDS, and PBS, resuspended in 50 μL
SDS-PAGE sample buffer, boiled, and subjected to SDS-PAGE
(7.5%). Labeled proteins were detected by phosphorimaging
20 (Molecular Dynamics). The results are presented in Figure 9.

6.9. Peptide Competition of GST-SH3 Affinity Precipit-ations of PI-3' Kinase From Cell Lysates

      It is possible to follow the precipitation of PI-3'
25 Kinase by Src from cell lysates in the presence or absence of
SH3-binding peptides. HeLa cells are lysed with detergent
and the protein mixtures are incubated for several hours with
the Src-GST fusion protein immobilized on glutathione-linked
Sepharose beads. After several hours of tumbling, the beads
30 are pelleted gently by low-speed centrifugation and the
supernatant is discarded. The beads are then resuspended
into a slurry in PBS-0.1% Tween 20, pelleted, and washed
several additional times. Finally, an SDS solution is added
to the sample, which is then boiled at 100 °C for 3 minutes.
35 Subsequently, the sample is centrifuged, and the supernatant
is loaded on a 10% polyacrylamide SDS gel for
electrophoresis. After the proteins hav been resolved, the

- 48 -

SUBSTITUTE SHEET (RULE 26)

gel is blotted to nitrocellulose or nylon (i.e., western blot). The filter is then probed with a PI-3' Kinase antibody (monoclonal and polyclonal antibodies are available from Upstate Biotechnology Incorporated, Lake Placid, NY) and

5   an enzyme-linked secondary antibody. The amount of PI-3' Kinase is then quantitated.

The ability of Src SH3 to bind PI-3' Kinase is examined for competability with exogenous peptides. Synthetic peptides corresponding to phage-displayed inserts and motifs

10 are added at the time that the lysate is incubated with the Src-GST fusion protein that has been immobilized on glutathione-linked sepharose beads. Ten-fold and one hundred-fold molar excess of peptides are used relative to SH3 proteins. The SH3 binding peptides block binding of the

15 PI-3' Kinase as detected on western blots while negative control peptides (unrelated peptide sequences) do not. The amount of competition is quantified and correlated with the amount of added SH3-domain binding peptides.

20   6.10.   *In Vivo* Association Of SH3-Binding Peptides
                   With SH3-Domains Of Proteins

To demonstrate association of the SH3-binding peptides with SH3-domains of proteins inside cells, the SH3-binding peptides are tagged and localized in cells. For

25 example, Bar-Sagi et al., in *Cell* (1993) 74:83-91, have shown that SH3-binding proteins localize to the cytoskeleton when expressed in cells. Thus, the SH3 domain-binding peptides of the present invention can serve as cellular targetting signals (e.g., to the cytoskeleton). Accordingly, the

30 peptides are tagged with biotin and, subsequently, injected into cells. Alternatively, one can transfect into cells a recombinant plasmid that expresses a fusion protein comprising of the SH3-binding peptide and the green fluorescent protein (GFP, Chalfie et al., in *Science* (1994)

35 263:802-805). The location of the biotinylated peptide or the GFP fusion protein is then assayed with FITC-labeled streptavidin in paraformaldehyde-fixed cells or by direct

- 49 -

SUBSTITUTE SHEET (RULE 26)

fluor scence in living cells, respectively. Localization of
the SH3-binding peptides to th cytoskeleton demonstrates
that the SH3-binding peptides can bind SH3-domain proteins *in
vivo*. In addition, focal adhesions, which are rich in Src,
5 are also sites of potential subcellular localization of SH3-
binding peptides.

Thus, NIH 3T3 fibroblasts were cultured *in vitro* on
glass coverslips coated with fibronectin. After two days of
growth at 37 °C, the cells were fixed for one hour at room
10 temperature in the presence of 2% paraformaldehyde (pH 7.5).
The coverslips were washed with PBS-0.1% Tween 20 several
times to remove the fixative. Next, the coverslips were
dipped into acetone (chilled at -20 °C) for approximately 20
seconds and allowed to air-dry. The coverslips were washed
15 again with PBS-0.1% Tween 20, containing BSA (1 mg/mL), and
incubated for 2 hours at room temperature with different
biotinylated peptides in PBS-0.1% Tween 20. The coverslips
were washed and then incubated with 1 mg/mL streptavidin-Cy3
(Jackson Immunoresearch Co., West Grove, PA) for 1 hour at
20 room temperature. Finally, the coverslips were washed in
PBS-0.1% Tween 20, mounted in a glycerol solution on a glass
slide, and viewed with a Nikon Optiphot epifluorescence
microscope and a 60x oil immersion lens.

The results are presented in Figure 11, in which panel A
25 displays cells stained with the conjugate biotin-spacer-
VLKRPLPIPPVTR (SEQ ID NO:64); panel B exhibits cells stained
with the conjugate, biotin-spacer-GILAPPVPPRNTR (SEQ ID
NO:63); panel C shows cells stained with the long consensus
peptide, biotin-spacer-RSTPRPLPPLPTTR (SEQ ID NO:67); and
30 panel D shows cells stained with the proline-rich vinculin
peptide conjugate, biotin-spacer-LAPPKPPLPEGEV (SEQ ID
NO:70). The "spacer" sequence is KSGSG (SEQ ID NO:103). As
shown in Figure 11, the panels in which SH3 domain-binding
peptides were used present a bright display of fluorescence
35 activity that is in sharp contrast to the relatively "dark"
featur s of panel D (non-SH3 domain binding vinculin
segment). These results d monstrate further the ability of

- 50 -

the SH3 domain-binding peptides of the present invention to
localize to protein targets (e.g., Src and Src-related
prot ins) within cells and provide an image thereof.

5   6.11.   *In Vivo Modulation Of Src In Oocytes With
              SH3-Binding Peptides*

When *Xenopus laevis* oocytes are injected with
mRNA encoding deregulated Src, there are *dramatic* cytological
and biochemical changes in the oocyte (Unger, T.F. and
Steele, R.E., in Mol. Cell. Biol. (1992) 12:5485-5498). The
applicants have obtained the plasmid for generating Src mRNA,
which is available from Dr. Robert Steele (University of
California at Irvine). Synthetic SH3-binding peptides are
injected into oocytes that have been previously injected with
Src mRNA. The state of the cytoskeleton is inspected
visually by observing the arrangement of cortical pigment
granules under a dissecting microscope. The state of
phosphorylation of several proteins is examined by western
blotting with an anti-phosphotryosine monoclonal antibody
(4G10; Upstate Biotechnology Incorporated), as described in
Unger and Steele, above.

6.12.   **Progesterone-induced *X. laevis* Oocyte
          Maturation**

Segments of adult ovary were removed
25 surgically and incubated in 0.1% collagenase type D
(Boehringer Mannheim, Indianapolis, IN) in Ca^{2+}-free OR2 (82.5
mM NaCl, 2.5 mM KCl, 1.0 mM $MgCl_2$, 1.0 mM Na_2HPO_4, 5.0 mM
HEPES, and 3.8 mM NaOH, pH 7.6). Oocytes were then washed 3-
5 times with OR2 containing 1.0 mM $CaCl_2$ and allowed to
30 recover in OR2 overnight at 18 °C. Stage VI oocytes were
injected with 40 nL of 100 mM peptide or water. After
injection, the oocytes were placed in OR2 with 2 mg/mL
progesterone (Sigma, St Louis, MO) and incubated at 20 °C.
Oocytes were scored at hourly time points for germinal
35 vesicle breakdown (GVBD).

- 51 -

Figure 10 presents the results of this experiment. As
shown by the graph, oocytes injected with the SH3 domain-
binding peptide VLKRPLPIPPVTR (SEQ ID NO:64) xhibit a faster
rate of progesterone-induced germinal vesicle breakdown
5 relative to oocytes that had been injected with water or with
the proline-rich vinculin peptide, LAPPKPPLPEGEV (SEQ ID
NO:70). These results parallel those obtained by Unger and
Steele, *supra*, who observed a gross alteration in the cortex
of oocytes that had been injected with RNA encoding a
10 deregulated Src variant ("active" Src) versus those injected
with RNA encoding the wild-type Src ("cellular" Src). Also,
as shown in Figure 3B of the Unger and Steele article,
oocytes injected with deregulated or active Src RNA matured
at a faster rate than oocytes injected with water or wild-
15 type Src RNA.

The present results obtained with Src SH3 domain-binding
peptides suggest that these peptides modulate the biochemical
activity of "cellular" Src; in particular, it is proposed
that at least some of the Src SH3 domain-binding peptides of
20 the present invention upregulate the biochemical activity of
"cellular" Src, which may be downregulated or inhibited in
its normal state. Hence, the administration of the SH3
domain-binding peptides of the present invention can
constitute a novel method of modulating the activity of Src
25 or Src-related proteins. Specifically, certain of these
peptides are able to activate Src-family proteins.

**6.13.   *In Vivo* Antagonism Of Src In Src Transformed
               Cells With SH3-Binding Peptides**

30
The coding regions for SH3-binding peptides
are cloned into vectors that direct their expression in
animal cells. A bipartite gene is constructed, encoding a
protein with *c-myc* epitope and SH3-binding peptide, which is
transcribed from a strong constitutive promoter (e.g., SV40,
35 CMV, HSV TK, calmodulin). The vector is introduced into
either normal or Src-transformed cells via transfection
(e.g., electroporation, calcium phosphate, liposomes, DEAE

- 52 -

SUBSTITUTE SHEET (RULE 26)

dextran). Transfected cells express the bipartite g ne transiently in culture. To create stable transformed cell lines, the vector carries a selectable marker (e.g., neomycin resistance) or transfection is performed in the presence of

5 excess plasmid carrying a selectable marker (e.g., neomycin resistance) and cells selected for the marker. Transfected cells are stained by immunofluorescence to detect expression of the bipartite protein. The hybridoma 9E10 secretes a monoclonal antibody that is highly specific for the *c-myc*

10 epitope (EQKLISEEDLN [SEQ ID NO:105]; see, Evan, G.A. et al., in Mol. Cell. Biol. (1985) 5:3610-3616). This antibody is used in immunofluorescence experiments to demonstrate that the bipartite protein is expressed inside the cells, and in some cases, localized to subcellular structures enriched in

15 SH3 domain bearing proteins.

There are several controls used in these experiments. First, cells are transfected with vectors that do not have the SH3-binding peptide coding region. Second, normal (non-transformed) cells are transfected. Third, cells transformed

20 by oncogenes other than Src are used in the transfection experiments. Fourth, cells are stained with other monoclonal antibodies that do not recognize the *c-myc* epitope.

Transfected cells are examined for any changes in cell shape, behavior, and metabolism as a consequence of

25 expressing the SH3 binding peptides. Cell shape is examined by phase contrast microscope at several times after transfection; in particular, the flatness of the cells, their adhesion to the substrate, and the degree of cell ruffling are monitored. Cell division rates, cell migration, and

30 contact inhibition are also observed over time. Finally, the amount of phosphorylated tyrosine in transfected cells is quantitated by phosphoaminoacid analysis and with an anti-phosphotryosine monoclonal antibody (4G10; Upstate Biotechnology Incorporated) in western blotting experiments.

35

SUBSTITUTE SHEET (RULE 26)

6.14 Pr paration f PXXP Bias d P ptid Librarie

The preparation and charact rization of preferred phage-displayed random peptide libraries have been described above in Sections 6.1 - 6.4.

5   Using procedures similar to those described in these sections, oligonucleotide inserts were contructed according to the schematic provided in FIG. 12. The inserts were then cloned into the mBAX vector, and the biased peptide librari s were expressed as described previously.

10   The mBAX vector was created in the Kay Laboratory by generating cloning sites in gene III of the M13mp18 vector (Messing, J. (1991). Cloning in M13 phage or how to use biology at its best. *Gene* 100, 3-12) in the manner of Fowlkes et al. (1992)(*Biotechniques* 13, 422-427). The mBAX

15 vector displays a peptide sequence at the N-terminus of the mature pIII protein that encodes the epitope for the mous monoclonal antibody 7E11 (see, FIG. 13); it includes the stop codon TAG in the coding region, which is suppressed in *E. coli* carrying suppressor tRNA gene mutations known as *supE* or

20 *supF*. There are no other stop codons in the mBAX genome. The mBAX vector also carries a segment of the alpha fragment of ß-galactosidase; bacterial cells expressing the omega fragment of ß-galactosidase convert the clear XGal substrat into an insoluble blue precipitate. The plaques appear blu .

25   Recombinant mBAX molecules can be distinguished easily from non-recombinant molecules due to the TAG codon in the *Xho*I - *Xba*I segment in gene III of mBAX. When recombinants are generated by replacing the *Xho* I - *Xba* I fragment with oligonucleotides encoding random peptides, the recombinants

30 can be grown in bacteria with (i.e., DH5αF') or without (i.e., JS5) suppressor tRNA mutant genes. On the other hand, the non-recombinant mBAX molecules fail to produce plaques on bacterial lawns where the bacteria (i.e., JS5) lack such suppressor genes. This is because in JS5, the TAG codon

35 serves as a stop codon to yield a truncated pIII molecule during translation; since pIII is an essential prot in

- 54 -

SUBSTITUTE SHEET (RULE 26)

component of viable M13 viral particles, no plaques will form.

6.14.1   Discussion

The use of second generation or biased peptide

5 libraries, which fix all or part of the RPLPPLP consensus motif for SH3 binding peptides and randomize flanking residues, has defined additional sequence residues exhibiting selective SH3 binding.

Tables 1-5, below, list the relevant amino acid

10 sequences obtained from the biased peptide library for each set of SH3 domain binding peptides. The underscored amino acid residues indicate the fixed positions. Also, indicated for each set of new binders is a "consensus" sequence, which seeks to include the additional features gleaned from the new

15 binding peptides. The symbol "ϕ" represents a hydrophobic residue.

TABLE 1   CORTACTIN SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.CORT.M1/2/3.PP	SSLLGPPVPPKPQTLFSFSR	107
PXXP.CORT.M4.PP	SRLGEFSKPPIPQKPTWMSR	108
PXXP.CORT.N2.PP	SRTERPPLPQRPDWLSYSSR	109
PXXP.CORT.N3.PP.INC	SREPDWLCPNCPLLLRSDSR	110
PXXP.CORT.01/2/3.PP	SSSSHNSRPPLPEKPSWLSR	111
PXXP.CORT.04.PP	SRLTPQSKPPLPPKPSAVSR	112
CONSENSUS	KPPϕPxKPxW R	113

TABLE 2   NCK SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.NCK.Q1/4.PP	SSLGVGWKPLPPMRTASLSR	114
PXXP.NCK.Q2/3.PP.INC	SSVGFADRPRPPLRVESLSR	115
PXXP.NCK.R1.PP.INC	SSAGILRPPEKPXRSFSLSR	116
PXXP.NCK.R2.PP	SSPYTGDVPIPPLRGASLSR	117
PXXP.NCK.R3.PP	SSLMGSWPPVPPLRSDSLSR	118
PXXP.NCK.R4.PP	SSIGEDTPPSPPTRRASLSR	119
PXXP.NCK.S1/4.PP	SRSLSEVSPKPPIRSVSLSR	120
PXXP.NCK.S2.PP.INC	SSVSEGYSPPLPPRSTSLSR	121
PXXP.NCK.S3.PP	SSSFTLAAPTPPTRSLSLSR	122
PXXP.NCK.T1.PP	SSPPYELPPRPPNRTVSLSR	123
PXXP.NCK.T2.PP	SRVVDGLAPPPPVRLSSLSR	124
PXXP.NCK.T3.PP.INC	SSLGYSGAPVPPHRxSSLSR	125
PXXP.NCK.T4.PP	SSISDYSRPPPPVRTLSLSR	126
CONSENSUS	φxxxxxPxPPφRSxSL T	127

BNSDOCID: <WO___9603649A1_IA>

TABLE 3   ABL SH3 BINDING PEPTIDES

		SEQ. ID NO.
PXXP.ABL.G1/2.PP	SRGPRWSPPPVPLPTSLDSR	128
PXXP.ABL.G3/4.PP	SSPPDYAAPAIPSSLWVDSR	129
PXXP.ABL.H1/3/4.PP	SSPPHWAPPAPPAMSPPISR	130
PXXP.ABL.H2.PP.INC	SSDRCWECPPWPAGGQRGSR	131
PXXP.ABL.I1/2/3.PP	SSPPKFSPPPPPYWQLHASR	132
PXXP.ABL.I4.PP	SSPPSFAPPAAPPRHSFGSR	133
PXXP.ABL.J1.PP	SSAPKKPAPPVPMMAHVMSR	134
PXXP.ABL.J2.PP.INC	SSPTYPPPPPPDTAKGASR	135
PXXP.ABL.J3.PP.INC	SSPPXXXPPPIPNSPQVLSR	136
PXXP.ABL.J4.PP	SSPPTWTPPKPPGWGVVFSR	137
PXXP.ABL.L1.PP	SSAPTWSPPALPNVAKYKSR	138
PXXP.ABL.L2/3.PP	SSIKGPRFPVPPVPLNGVSR	139
PXXP.ABL.L4.PP	SSPPAWSPPHRPVAFGSTSR	140
CONSENSUS	PPxWxPPPϕP	141

TABLE 4   PLCG SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.PLCG.P1.PP	SSMKVHNFPLPPLPSYETSR	142
PXXP.PLCG.P2.PP	SRVPPLVAPRPPSTLNSLSR	143
PXXP.PLCG.PE.PP.INC	SSLYWQHGPDPPVGAPQLSR	144
PXXP.PLCG.P4.PP	SSHPLNSWPGGPFRHNLSSR	145

- 57 -

SUBSTITUTE SHEET (RULE 26)

TABLE 5   SRC SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.SRC.A1.PP	SSRALRVRPLPPVPGTSLSR	146
PXXP.SRC.A2.PP	SSFRALPLPPTPDNPFAGSR	147
PXXP.SRC.A3.PP	SRDAPGSLPFRPLPPVPTSR	148
PXXP.SRC.A4.PP	SSISQRALPPLPLMSDPASR	149
PXXP.SRC.B1.PP	SSPAYRPLPRLPDLSVIYSR	150
PXXP.SRC.B2/3/PP	SSFINRRLPALPPDNSLLSR	151
PXXP.SRC.B4.PP	SRLTGRPLPALPPPFSDFSR	152
PXXP.SRC.C1.PP	SRMKDRVLPPIPTVESAVSR	153
PXXP.SRC.C2.PP.INC	SSLYSAIAPDPPPRNSSSSR	154
PXXP.SRC.C3.PP	SSLASRPLPLLPNSAPGQSR	155
PXXP.SRC.D1.PP	SSLTSRPLPDIPVRPSKSSR	156
PXXP.SRC.D2.PP.INC	SSLKWRALPPLPETDTPYSR	157
PXXP.SRC.D3.PP	SSNTNRLPPPTPDGLDVRSR	158
PXXP.SRC.D4.PP	SSLQSRPLPLPPQSSYPISR	159
CONSENSUS	RPLPPLP	160

It should be apparent to one of ordinary skill that many other embodiments of the present invention can be contemplated beyond the preferred embodiments described above but which other embodiments nevertheless fall within the scope and spirit of the present invention. Hence, the present invention should not be construed to be limited to the preferred embodiments described herein, which serve only to illustrate the present invention, but only by the claims that follow.

Also, numerous references are cited throughout the specification. The complete disclosures of these references are incorporated by reference herein.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: SPARKS, ANDREW B.
                  KAY, BRIAN K.
                  THORN, JUDITH M.
                  QUILLIAM, LAWRENCE A.
                  DER, CHANNING J.

    (ii) TITLE OF INVENTION: Src SH3 BINDING PEPTIDES AND METHODS OF
          ISOLATING AND USING SAME

    (iii) NUMBER OF SEQUENCES: 106

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Pennie & Edmonds
        (B) STREET: 1155 Avenue of the Americas
        (C) CITY: New York
        (D) STATE: New York
        (E) COUNTRY: U.S.A.
        (F) ZIP: 10036-2711

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: S. Leslie Misrock
        (B) REGISTRATION NUMBER: 18,872
        (C) REFERENCE/DOCKET NUMBER: 1101-192

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (212) 790-9090
        (B) TELEFAX: (212) 869-9741/8864
        (C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8
        (D) OTHER INFORMATION: /note= "hydrophobic residue (such
          as Pro or Leu)"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Xaa Pro Xaa Xaa Pro Pro Pro Xaa Xaa Pro
1   5   10

- 59 -

SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 1..3
        (D) OTHER INFORMATION: /note= "Xaa = any residue other
            than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /note= "X = any residue other
            than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8..9
        (D) OTHER INFORMATION: /note= "X = any residue other
            than Cys"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

    Xaa Xaa Xaa Pro Pro Xaa Pro Xaa Xaa
    1   5

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "X tends to be hydrophobic"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8..9
        (D) OTHER INFORMATION: /note= "X tends to be hydrophobic"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

    Arg Xaa Leu Pro Pro Arg Pro Xaa Xaa
    1   5

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

- 60 -

SUBSTITUTE SHEET (RULE 26)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Lys Leu Pro Pro Arg Pro Arg Arg
1   5

(2) INFORMATION FOR SEQ ID NO:5:

5

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 10 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

10

Val Pro Pro Pro Val Pro Pro Arg Arg Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 7 amino acids
       (B) TYPE: amino acid

15

       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 2
       (D) OTHER INFORMATION: /note= "tends to be Pro"

20

    (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 3
       (D) OTHER INFORMATION: /note= "hydrophobic residue"

    (ix) FEATURE:
       (A) NAME/KEY: Modified-site
       (B) LOCATION: 5

25

       (D) OTHER INFORMATION: /note= "tends to be Pro"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Xaa Xaa Xaa Pro Xaa Xaa Pro
1   5

30 (2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 9 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

35

    (ix) FEATURE:
       (A) NAME/KEY: Modifi d-site
       (B) LOCATION: 2
       (D) OTHER INFORMATION: /note= "X = any residue other
           than Cys"

- 61 -

SUBSTITUTE SHEET (RULE 26)

```
        (ix) FEATURE:
             (A) NAME/KEY: Modifi d-site
             (B) LOCATION: 8..9
             (D) OTHER INFORMATION: /note= "X = any residu  oth r
                      than Cys"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
```

5
```
    Arg Xaa Leu Pro Pro Arg Pro Xaa Xaa
    1           5
```

```
    (2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 9 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown
```
10

```
        (ii) MOLECULE TYPE: peptide

        (ix) FEATURE:
             (A) NAME/KEY: Modified-site
             (B) LOCATION: 9
             (D) OTHER INFORMATION: /note= "hydrophobic residue"
```

15
```
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

        Arg Xaa Leu Pro Pro Leu Pro Arg Phe
        1           5
```

```
    (2) INFORMATION FOR SEQ ID NO:9:
```
20
```
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 7 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
```
25
```
        Arg Pro Leu Pro Pro Leu Pro
        1           5
```

```
    (2) INFORMATION FOR SEQ ID NO:10:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 9 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown
```
30
```
        (ii) MOLECULE TYPE: peptide

        (ix) FEATURE:
             (A) NAME/KEY: Modified-site
             (B) LOCATION: 2
             (D) OTHER INFORMATION: /note= "any residue other than Cys"
```

35
```
        (ix) FEATURE:
             (A) NAME/KEY: Modifi d-sit
             (B) LOCATI N: 5..6
             (D) OTHER INFORMATION: /note= "any hydrophobic residue"
```

- 62 -

SUBSTITUTE SHEET (RULE 26)

```
       (ix) FEATURE:
            (A) NAME/KEY: Modified-site
            (B) LOCATION: 9
            (D) OTHER INFORMATION: /note= "any hydrophilic residue
                     except Cys"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

5
       Arg Xaa Leu Pro Xaa Xaa Pro Xaa Xaa
       1               5


    (2) INFORMATION FOR SEQ ID NO:11:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 9 amino acids
10         (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: peptide

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

       Arg Pro Leu Pro Pro Leu Pro Thr Ser
       1               5
15

    (2) INFORMATION FOR SEQ ID NO:12:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 4 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: linear

20    (ii) MOLECULE TYPE: peptide

       (v) FRAGMENT TYPE:

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

       Xaa Xaa Xaa Xaa
       1

25
   (2) INFORMATION FOR SEQ ID NO:13:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 5 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: peptide

30     (v) FRAGMENT TYPE:

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

       Xaa Xaa Xaa Xaa Xaa
       1               5


35 (2) INFORMATION FOR SEQ ID NO:14:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 4 amino acids
           (B) TYPE: amino acid
```

- 63 -

SUBSTITUTE SHEET (RULE 26)

                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: peptide

            (v) FRAGMENT TYPE:

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

5
        Xaa Xaa Xaa Xaa
        1

    (2) INFORMATION FOR SEQ ID NO:15:

            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 13 amino acids
                (B) TYPE: amino acid
10   (D) TOPOLOGY: unknown

            (ii) MOLECULE TYPE: peptide

            (ix) FEATURE:
                (A) NAME/KEY: Modified-site
                (B) LOCATION: 1..3
                (D) OTHER INFORMATION: /note= "any residue other than Cys"

15   (ix) FEATURE:
                (A) NAME/KEY: Modified-site
                (B) LOCATION: 5
                (D) OTHER INFORMATION: /note= "any residue other than Cys"

            (ix) FEATURE:
                (A) NAME/KEY: Modified-site
                (B) LOCATION: 8..9
20   (D) OTHER INFORMATION: /note= "any hydrophobic residue"

            (ix) FEATURE:
                (A) NAME/KEY: Modified-site
                (B) LOCATION: 11..13
                (D) OTHER INFORMATION: /note= "any residue other than Cys"

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

25
        Xaa Xaa Xaa Arg Xaa Leu Pro Xaa Xaa Pro Xaa Xaa Xaa
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:16:

            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 42 amino acids
30   (B) TYPE: amino acid
                (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: peptide

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

        Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        1   5   10   15
35
        Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                        20   25   30

        Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg

                            - 64 -

                    SUBSTITUTE SHEET (RULE 26)

35   40

(2) INFORMATION FOR SEQ ID NO:17:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 69 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGCTCGAGNN NBNNBNNBNN BNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNNBNNBNNB   60

NNBCCAGGT   69

(2) INFORMATION FOR SEQ ID NO:18:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 68 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGTCCANNVN NVNNVNNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVNN VNNVNNVNNV   60

AGATCTGG   68

(2) INFORMATION FOR SEQ ID NO:19:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 66 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCGAGNNNBN NBNNBNNBNN BNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNNBNNBNNB   60

CCAGGT   66

(2) INFORMATION FOR SEQ ID NO:20:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 65 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

- 65 -

GGTCCANNVN NVNNVNNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVNN VNNVNNVNNV   60

AGATC   65

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 46 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE:

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 2
    (D) OTHER INFORMATION: /note= "Arg or Ser"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1   5   10   15

Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20   25   30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Pro Ser Arg Thr
        35   40   45

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 49 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Ser His Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1   5   10   15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20   25   30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Pro Ser Arg
        35   40   45

Thr

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 27 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- 66 -

SUBSTITUTE SHEET (RULE 26)

                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 2
                    (D) OTHER INFORMATION: /note= "Ser or Thr"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 13
     5   (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 15
                    (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

    10   Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa
         1   5   10   15

         Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg
                         20   25

       (2) INFORMATION FOR SEQ ID NO:24:

            (i) SEQUENCE CHARACTERISTICS:
    15   (A) LENGTH: 48 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: both
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: DNA (genomic)

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

    20 TTTTGTCGAC NNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNGCGGTG   48

       (2) INFORMATION FOR SEQ ID NO:25:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 48 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: both
    25   (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: DNA (genomic)

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

       CGCCACNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVTG ATCATTTT   48

    30 (2) INFORMATION FOR SEQ ID NO:26:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 43 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: double
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: DNA (genomic)

    35   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

       CGCCACNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVTG ATC   43

                                - 67 -

                        SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 43 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

5

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCGACNNNBN NBNNBNNBNN BNNBNNBNNB NNBNNBNGCG GTG   43

(2) INFORMATION FOR SEQ ID NO:28:

10

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (v) FRAGMENT TYPE:

15

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "Ser or Thr"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 13

20

        (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 15
        (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

25

Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa
1   5   10   15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg Pro Ser Arg Thr
        20   25   30

(2) INFORMATION FOR SEQ ID NO:29:

30

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 34 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site

35

        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "Ser or Thr"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site

- 68 -

SUBSTITUTE SHEET (RULE 26)

                        (B) LOCATION: 16
                        (D) OTHER INFORMATION: /not = "Ser, Arg, Gly, Cys or Trp"

                (ix) FEATURE:
                        (A) NAME/KEY: Modified-site
                        (B) LOCATION: 18
                        (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

                (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

        Ser His Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        1   5   10   15

        Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg Pro Ser
                        20   25   30

        Arg Thr

    (2) INFORMATION FOR SEQ ID NO:30:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 8 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: peptide

            (v) FRAGMENT TYPE: internal

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

        Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        1   5

    (2) INFORMATION FOR SEQ ID NO:31:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 58 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: both
                    (D) TOPOLOGY: unknown

            (ii) MOLECULE TYPE: DNA (genomic)

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

    TGACGYCTCG AGTTGTNNKN NKNNKNNKNN KNNKNNKNNK TGTGGATCTA GAAGGATC   58

    (2) INFORMATION FOR SEQ ID NO:32:

            (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 15 base pairs
                    (B) TYPE: nucleic acid
                    (C) STRANDEDNESS: both
                    (D) TOPOLOGY: unknown

            (ii) MOLECULE TYPE: DNA (genomic)

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

                                    - 69 -

CCTAGATCTT CCTAG   15

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 40 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TCGAGTTGTN NKNNKNNKNN KNNKNNKNNK NNKTGTGGAT   40

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 40 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CAACANNMNN MNNMNNMNNM NNMNNMNNMA CACCTAGATC   40

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (v) FRAGMENT TYPE:

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ser Ser Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Gly Ser Arg Pro
1   5   10   15

Ser Arg Thr

(2) INFORMATION FOR SEQ ID NO:36:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Ser Ser Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Gly Ser Arg Ser
1   5   10   15

- 70 -

SUBSTITUTE SHEET (RULE 26)

```
        Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Thr Arg
                 20                      25
```

      (2) INFORMATION FOR SEQ ID NO:37:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 80 base pairs
              (B) TYPE: nucleic acid
              (C) STRANDEDNESS: double
              (D) TOPOLOGY: unknown

         (ii) MOLECULE TYPE: DNA (genomic)

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```
TCGAGTTGTN NKNNKNNKNN KNNKNNKNNK NNKTGTGGAT CTAGATCCAC AVNNVNNVNN    60

VNNVNNVNNV NNVNNACAAC                                                80
```

      (2) INFORMATION FOR SEQ ID NO:38:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 45 amino acids
              (B) TYPE: amino acid
              (D) TOPOLOGY: unknown

         (ii) MOLECULE TYPE: peptide

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
        Ser Ser Phe Asp Gln Gln Asp Trp Asp Tyr Ser Ile Ala Glu Lys Met
        1               5                   10                  15

        His Pro Ile Arg Pro Gly Phe Arg Glu Leu Pro Pro Leu Pro Pro Ser
                     20                  25                  30

        Arg Ala Ser Phe Gly Gly Gly Ala Ser Arg Pro Ser Arg
                     35                  40                  45
```

      (2) INFORMATION FOR SEQ ID NO:39:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 30 amino acids
              (B) TYPE: amino acid
              (D) TOPOLOGY: unknown

         (ii) MOLECULE TYPE: peptide

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
        Ser Thr Asn Val Trp Val Thr Gly Ser Val Ile Ala Arg Gly Ala Gln
        1               5                   10                  15

        Ser Arg Pro Leu Pro Ile Pro Pro Glu Thr Arg Pro Ser Arg
                     20                  25                  30
```

      (2) INFORMATION FOR SEQ ID NO:40:

          (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 30 amino acids
              (B) TYPE: amino acid

                          - 71 -

                    SUBSTITUTE SHEET (RULE 26)

                    (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

          Ser Thr Ala Pro Trp Gly Leu Arg Val Ala His Glu Gly Gly Val Leu
     5   1   5   10   15

          Lys Arg Pro Leu Pro Ile Pro Pro Val Thr Arg Pro Ser Arg
                    20   25   30

     (2) INFORMATION FOR SEQ ID NO:41:

          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 45 amino acids
    10   (B) TYPE: amino acid
               (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

          Ser Ser Ser Gly Tyr Val Val Pro Lys Arg Leu Gly Asp Met Arg Glu
    15   1   5   10   15

          Tyr Asn Ala His Pro Gly Leu His Val Pro Pro Asn Ser Pro Leu Pro
                    20   25   30

          Pro Leu Pro Thr His Leu Gln Ser Ser Arg Pro Ser Arg
                    35   40   45

    20 (2) INFORMATION FOR SEQ ID NO:42:

          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 44 amino acids
               (B) TYPE: amino acid
               (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide

    25   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

          Ser Ser Arg Gly Glu Gly Asn Asn Ile Ile Ser Ser Arg Pro Phe Leu
          1   5   10   15

          Ser Asn Ser Asp Pro Gly Val Ser Asn Lys Leu Thr Gly Arg Gly Pro
                    20   25   30

          Leu Pro Pro Leu Pro Asn Asp Ser Arg Pro Ser Arg
    30   35   40

     (2) INFORMATION FOR SEQ ID NO:43:

          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 31 amino acids
               (B) TYPE: amino acid
               (D) TOPOLOGY: unknown
    35
          (ii) MOLECULE TYPE: peptid

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

                              - 72 -

                    SUBSTITUTE SHEET (RULE 26)

```
Ser Thr Ala Val Ser Phe Arg Phe Met Pro Gly Gly Gly Gly Ala Phe
1               5               10              15
```

```
Tyr Ser Thr Arg Pro Val Pro Pro Ile Thr Arg Pro Ser Arg Thr
            20              25              30
```

5 (2) INFORMATION FOR SEQ ID NO:44:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

       (ii) MOLECULE TYPE: peptide

10

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```
Ser Thr Ala His Ser Leu Trp Asp Trp Gly Thr Phe Ser Gly Val Ser
1               5               10              15
```

```
His Lys Ser Arg Leu Pro Pro Leu Pro Thr Arg Pro Ser Arg Thr
            20              25              30
```

15

    (2) INFORMATION FOR SEQ ID NO:45:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 25 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

20

       (ii) MOLECULE TYPE: peptide

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Pro Gly Tyr Ala Arg Ile Val Ser Tyr Arg Phe Arg Ala Leu Pro Ser
1               5               10                      15
```

```
Pro Pro Ser Ala Ser Arg Pro Ser Arg
            20              25
```

25

    (2) INFORMATION FOR SEQ ID NO:46:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

30

       (ii) MOLECULE TYPE: peptide

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
Ser Thr Asn Asp Val Asp Trp Met His Met Trp Asn Ser Gly Gly Pro
1               5               10              15
```

```
His Arg Arg Leu Pro Pro Thr Pro Ala Thr Arg Pro Ser Arg
            20              25              30
```

35

    (2) INFORMATION FOR SEQ ID NO:47:

- 73 -

SUBSTITUTE SHEET (RULE 26)

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 45 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: peptide

5   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

      Ser Ser Asp Asn Trp Ala Arg Arg Val His Ala Ser Glu Leu Ile Tyr
      1   5   10   15

      Thr Asp Leu Ser Pro Gly Ile Leu Leu Ala Gln Arg Gln Leu Pro Pro
                      20   25   30

      Thr Pro Gly Arg Asp Pro Ser His Ser Arg Pro Ser Arg
                      35   40   45

   (2) INFORMATION FOR SEQ ID NO:48:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 45 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

15
      (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

      Ser Ser Glu Ser Pro Leu Met Tyr Asn Arg Val Gly Ala Leu Gln Ser
      1   5   10   15

      Leu Thr Ser Val Pro Gly Ser Met Met His Phe Ala Leu Gln Arg Arg
                      20   25   30

      Leu Pro Arg Thr Pro Pro Pro Ala Ser Arg Pro Ser Arg
                      35   40   45

   (2) INFORMATION FOR SEQ ID NO:49:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 35 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

      Ser Thr Arg Trp Ser His Ser Trp Pro Gly Tyr Val Gly Gly Ala Asn
      1   5   10   15

      Pro Ser Pro Ala Thr Arg Pro Leu Pro Thr Arg Pro Ser Arg Thr Val
                      20   25   30

      Glu Ser Cys
                35

35 (2) INFORMATION FOR SEQ ID NO:50:

       (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 50 amino acids
            (B) TYPE: amino acid

                          - 74 -

                 SUBSTITUTE SHEET (RULE 26)

        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

        Ser Arg Tyr Asn Asp Leu Gly Thr Arg Pro Val Ser Glu Val Ile Lys
        1   5   10   15

        Tyr Asp Tyr Phe Pro Gly Tyr Ser Gln His Val Ile Thr Pro Asp Gly
                    20   25   30

        Ser Tyr Ser Thr Arg Pro Leu Pro Ser Arg Pro Ser Arg Thr Val Glu
                35   40   45

        Ser Cys
        50

    (2) INFORMATION FOR SEQ ID NO:51:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 24 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

        Pro Gly Arg Leu Leu Leu Pro Ser Glu Pro Arg Thr Phe Tyr Asn Tyr
        1   5   10   15

        Gly His Asp Ser Arg Pro Ser Arg
                    20

    (2) INFORMATION FOR SEQ ID NO:52:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 30 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

        Ser Thr Met Tyr Gly Val Ser Trp Leu Ser Ser Gly Ser Gly Gly Ile
        1   5   10   15

        Leu Ala Pro Pro Val Pro Pro Arg Asn Thr Arg Pro Ser Arg
                    20   25   30

    (2) INFORMATION FOR SEQ ID NO:53:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

                            - 75 -

                    SUBSTITUTE SHEET (RULE 26)

```
Ser Ser Cys Thr Glu Lys Thr Val Ser Gly Trp Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Pro Leu Pro Il  Leu Pro Arg Thr Thr Arg Pro Ser Arg
            20              25                  30
```

5 (2) INFORMATION FOR SEQ ID NO:54:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 31 amino acids
         (B) TYPE: amino acid
         (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

10   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```
Ser Ser Cys Met Leu Pro Thr Asp Gly Trp Gln Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg Pro Ser Arg
            20              25                  30
```

15 (2) INFORMATION FOR SEQ ID NO:55:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 31 amino acids
         (B) TYPE: amino acid
         (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

20   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```
Ser Ser Cys Asp Gly Thr Gln Phe Arg Leu Asn Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Asn Arg Pro Leu Pro Met Ile Pro Thr Thr Arg Pro Ser Arg
            20              25                  30
```

25 (2) INFORMATION FOR SEQ ID NO:56:

     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 31 amino acids
         (B) TYPE: amino acid
         (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

30   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```
Ser Ser Cys Met Gln Gly Gln Ala Gly Leu Lys Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr Arg Pro Ser Arg
            20              25                  30
```

35

     (2) INFORMATION FOR SEQ ID NO:57:

- 76 -

SUBSTITUTE SHEET (RULE 26)

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 31 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

5   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

    Ser Ser Cys Tyr Arg Glu Lys Asp Thr Trp Gly Cys Gly Ser Arg Ser
    1   5   10   15

    Thr Ser Arg Pro Leu Pro Ser Leu Pro Thr Thr Arg Pro Ser Arg
                20   25   30

10 (2) INFORMATION FOR SEQ ID NO:58:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 32 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

    Ser Ser Cys Leu Phe Glu Gln Gly Ala Gly Thr Cys Gly Ser Arg Ser
    1   5   10   15

    Thr Arg Ser Leu Pro Pro Leu Pro Pro Thr Thr Arg Pro Ser Ser Arg
                20   25   30

20 (2) INFORMATION FOR SEQ ID NO:59:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 44 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

    Ser Ser Cys Asp Thr Gly Arg Ile Ala Pro Gly Cys Gly Ser Arg Ser
    1   5   10   15

    Thr Pro Arg Pro Leu Pro Leu Ile Pro Thr Thr Pro Arg Ser Thr Asn
                20   25   30

    Leu Asn Leu Thr Ser Thr Thr Thr Arg Pro Ser Arg
30   35   40

    (2) INFORMATION FOR SEQ ID NO:60:

        (i) SEQUENCE CHARACTERISTICS:
35   (A) LENGTH: 31 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

                            - 77 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
Ser Ser Cys Gly Leu Asp Asn Ala Ala Lys Thr Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Pro Leu Pro Pro Thr Pro Leu Thr Thr Arg Pro Ser Arg
            20                  25                  30
```

5

(2) INFORMATION FOR SEQ ID NO:61:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

10   (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Ser Ser Cys Ser Arg Ala His Glu Thr Glu Met Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Pro Gln Pro Pro Pro Ile Thr Thr Arg Pro Ser Arg
            20                  25                  30
```

15

(2) INFORMATION FOR SEQ ID NO:62:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

20   (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```
Arg Ser Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:63:

25

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

30
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```
Gly Ile Leu Ala Pro Pro Val Pro Pro Arg Asn Thr Arg
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:64:

35   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

- 78 -

SUBSTITUTE SHEET (RULE 26)

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Val L u Lys Arg Pro Leu Pro Ile Pro Pro Val Thr Arg
1   5   10

5   (2) INFORMATION FOR SEQ ID NO:65:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

10   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Gly Pro His Arg Arg Leu Pro Pro Thr Pro Ala Thr Arg
1   5   10

    (2) INFORMATION FOR SEQ ID NO:66:

15   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

20   Ala Asn Pro Ser Pro Ala Thr Arg Pro Leu Pro Thr Arg
1   5   10

    (2) INFORMATION FOR SEQ ID NO:67:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 14 amino acids
       (B) TYPE: amino acid
25   (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Arg Ser Thr Pro Arg Pro Leu Pro Pro Leu Pro Thr Thr Arg
30   1   5   10

    (2) INFORMATION FOR SEQ ID NO:68:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 19 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

35   (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

- 79 -

SUBSTITUTE SHEET (RULE 26)

```
Gly Thr Val Glu Pro Val Pro Pro Pr  Val Pro Pro Arg Arg Arg Pro
1               5               10          15

Glu Ser Ala
```

(2) INFORMATION FOR SEQ ID NO:69:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 17 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```
Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro Arg Pro Leu Pro
1               5               10          15

Val
```

(2) INFORMATION FOR SEQ ID NO:70:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 13 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```
Leu Ala Pro Pro Lys Pro Pro Leu Pro Glu Gly Glu Val
1               5               10
```

(2) INFORMATION FOR SEQ ID NO:71:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 7 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 2
          (D) OTHER INFORMATION: /note= "any residue other than Cys"

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 5..6
          (D) OTHER INFORMATION: /note= "any hydrophobic residue"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```
Arg Xaa Leu Pro Xaa Xaa Pro
1               5
```

(2) INFORMATION FOR SEQ ID NO:72:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 13 amino acids

- 80 -

SUBSTITUTE SHEET (RULE 26)

```
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptid

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

5      Pro Gly Phe Arg Glu Leu Pro Pro Leu Pro Pro Ser Arg
       1               5                   10


    (2) INFORMATION FOR SEQ ID NO:73:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 13 amino acids
                    (B) TYPE: amino acid
10                  (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

       Ala Gln Ser Arg Pro Leu Pro Ile Pro Pro Glu Thr Arg
       1               5                   10

15
    (2) INFORMATION FOR SEQ ID NO:74:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 13 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide

20        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

       Pro Pro Asn Ser Pro Leu Pro Pro Leu Pro Thr His Leu
       1               5                   10



    (2) INFORMATION FOR SEQ ID NO:75:
25
          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 13 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

30     Thr Gly Arg Gly Pro Leu Pro Pro Leu Pro Asn Asp Ser
       1               5                   10


    (2) INFORMATION FOR SEQ ID NO:76:

          (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 13 amino acids
35                  (B) TYPE: amino acid
                    (D) TOPOLOGY: unknown

          (ii) MOLECULE TYPE: peptide
```

- 81 -

SUBSTITUTE SHEET (RULE 26)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Tyr S r Thr Arg Pro Val Pro Pro Ile Thr Arg Pr Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:77:

5

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

10

Ser His Lys Ser Arg Leu Pro Pro Leu Pro Thr Arg Pro
1   5   10

(2) INFORMATION FOR SEQ ID NO:78:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
15   (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Tyr Arg Phe Arg Ala Leu Pro Ser Pro Pro Ser Ala Ser
1   5   10

20

(2) INFORMATION FOR SEQ ID NO:79:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

25   (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Leu Ala Gln Arg Gln Leu Pro Pro Thr Pro Gly Arg Asp
1   5   10

(2) INFORMATION FOR SEQ ID NO:80:

30

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

35

Ala L u Gln Arg Arg L u Pr Arg Thr Pro Pro Pro Ala
1   5   10

- 82 -

SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:81:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 13 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: unknown

5 (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

   Pro Ala Thr Arg Pro Leu Pro Thr Arg Pro Ser Arg Thr
   1   5   10

   (2) INFORMATION FOR SEQ ID NO:82:
10
       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 13 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

15 Tyr Ser Thr Arg Pro Leu Pro Ser Arg Pro Ser Arg Thr
   1   5   10

   (2) INFORMATION FOR SEQ ID NO:83:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 13 amino acids
20   (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

   Xaa Pro Gly Arg Ile Leu Leu Leu Pro Ser Glu Pro Arg
25 1   5   10

   (2) INFORMATION FOR SEQ ID NO:84:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 13 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

30 (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

   Ser Gly Gly Ile Leu Ala Pro Pro Val Pro Pro Arg Asn
   1   5   10

35 (2) INFORMATION FOR SEQ ID NO:85:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 13 amino acids
           (B) TYPE: amino acid

- 83 -

SUBSTITUTE SHEET (RULE 26)

            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptid

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

        Arg Ser Thr Arg Pro Leu Pro Ile Leu Pro Arg Thr Thr
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:86:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

        Ser Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:87:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

        Ser Thr Asn Arg Pro Leu Pro Met Ile Pro Thr Thr Arg
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:88:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

        Arg Ser Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:89:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

                        - 84 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Ser Thr S r Arg Pro Leu Pro Ser Leu Pro Thr Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:90:

5

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

10

Arg Ser Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr
1   5   10

(2) INFORMATION FOR SEQ ID NO:91:

15

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

20

Arg Ser Thr Arg Gln Leu Pro Ile Pro Pro Thr Thr Thr
1   5   10

(2) INFORMATION FOR SEQ ID NO:92:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids

25

      (B) TYPE: amino acid
      (D) TOPOLOGY: unknown

(ii) MOLECULE TYPF: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

Ser Thr Pro Arg Pro Leu Pro Leu Ile Pro Thr Thr Pro
1   5   10

30

(2) INFORMATION FOR SEQ ID NO:93:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: unknown

35   (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

Arg Ser Thr Arg Pro Leu Pro Pr Thr Pr Leu Thr Thr

- 85 -

1   5   10

(2) INFORMATION FOR SEQ ID NO:94:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Arg Ser Thr Arg Pro Gln Pro Pro Pro Ile Thr Thr
1   5   10

(2) INFORMATION FOR SEQ ID NO:95:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Arg Pro Leu Pro Met Leu Pro
1   5

(2) INFORMATION FOR SEQ ID NO:96:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "Pro or Met"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

Arg Pro Leu Pro Xaa Leu Pro
1   5

(2) INFORMATION FOR SEQ ID NO:97:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

- 86 -

SUBSTITUTE SHEET (RULE 26)

Arg Ser Thr Pro
1

(2) INFORMATION FOR SEQ ID NO:98:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Arg Ser Thr Pro Ala Pro Pro Val Pro Pro Arg Thr Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:99:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

AGCGTAACGA TCTCCCG   17

(2) INFORMATION FOR SEQ ID NO:100:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TTCACCTCGA AAGCAAGCTG   20

(2) INFORMATION FOR SEQ ID NO:101:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CCTCATAGTT AGCGTAACG   19

- 87 -

SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:102:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

5   (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

AGCGTAACGA TCTAAA   16

10 (2) INFORMATION FOR SEQ ID NO:103:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

15   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

Lys Ser Gly Ser Gly
1   5

(2) INFORMATION FOR SEQ ID NO:104:

20   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

25   Gly Ser Gly Ser
1

(2) INFORMATION FOR SEQ ID NO:105:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 11 amino acids
        (B) TYPE: amino acid
30   (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
1   5   10
35

(2) INFORMATION FOR SEQ ID NO:106:

    (i) SEQUENCE CHARACTERISTICS:

- 88 -

SUBSTITUTE SHEET (RULE 26)

          (A) LENGTH: 31 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

     Ser Ser Cys Asp His Thr Leu Gly Leu Gly Trp Cys Gly Ser Arg Ser
     1   5   10   15

     Thr Arg Gln Leu Pro Ile Pro Pro Thr Thr Thr Arg Pro Ser Arg
               20   25   30

WHAT IS CLAIMED IS:

1.   A peptide having at 1 ast nine and up to forty-five amino acid residues, including an amino acid sequence of the formula, R-2-L-P-5-6-P-8-9 (SEQ ID NO:10), positioned anywhere along the peptide, in which each number represents an amino acid residue, such that 2 represents any amino acid residue except cysteine, 5 and 6 each represents a hydrophobic amino acid residue, 8 represents any amino acid residue except cysteine, and 9 represents a hydrophilic amino acid residue except cysteine, each letter being the standard one-letter symbol for the corresponding amino acid, said peptide exhibiting a binding affinity for the SH3 domain of Src, provided that said peptide is not R-P-L-P-P-L-P-T-S (SEQ ID NO:11).

2.   The peptide of claim 1 in which 2 is a P, R, A, L, Q, E or S.

3.   The peptide of claim 1 in which 5 is a P, M, I or L.

4.   The peptide of claim 1 in which 6 is a P, L, I or V.

5.   The peptide of claim 1 in which 8 is a T, R, P, I, N, E, V, S, A, G or L.

6.   The peptide of claim 1 in which 9 is a T, R, S, H or D.

7.   The peptide of claim 1 which further comprises a C-terminal-flanking amino acid sequence of the formula 10, 10-11, 10-11-12, 10-11-12-13 (SEQ ID NO:12) or 10-11-12-13-14 (SEQ ID NO:13), in which each number represents any amino acid residue except cysteine, such that 10 is bound to 9 by a peptide bond.

8.   The peptide of claim 7 in which 10 is T, R, L, S, D, P, A or N.

9.   The peptide of claim 7 in which 11 is R, P, A, Q, S or T.

10.   The peptide of claim 7 in which 12 is P, S, R or T.

11.   The peptide of claim 7 in which 13 is P, S, R, F, H or T.

SUBSTITUTE SHEET (RULE 26)

12.   Th peptide of claim 7 in which 14 is S, R, G or T.

13.   The peptide of claim 1 which further comprises an N-t rminal-flanking amino acid sequence of the formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in which each number represents any amino acid residue except cysteine, such that 1' is bound to R by a peptide bond.

14.   The peptide of claim 13 in which 1' is T, P, S, N, F, W, K, H, Q or G.

15.   The peptide of claim 13 in which 2' is S, T, G, P, R, Q, L, A or H.

16.   The peptide of claim 13 in which 3' is R, S, P, G, A, V, Y or L.

17.   The peptide of claim 13 in which 4' is R, S, V, T, G, L or F.

18.   A peptide having at least thirteen and up to forty-five amino acid residues, including an amino acid sequence of the formula, 3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned anywhere along the peptide, in which each number represents an amino acid residue, such that 3', 2', 1', 2, 8, and 10 each represents any amino acid residue except cysteine, 5 and 6 each represents a hydrophobic amino acid residue, and 9 represents a hydrophilic amino acid residue except cysteine, each letter being the standard one-letter symbol for the corresponding amino acid, said peptide exhibiting a binding affinity for the SH3 domain of Src.

19.   The peptide of claim 18 in which 5 is a P or M.

20.   The peptide of claim 18 in which 1' is T, P, S or N.

21.   The peptide of claim 18 in which 2' is S or T.

22.   The peptide of claim 18 in which 3' is R or S.

23.   The peptide of claim 18 in which 10 is T or R.

24.   The peptide of claim 1 the binding affinity of which is at least three-fold greater than that exhibited by the peptide RPLPPLP for the SH3 domain of Src.

25.   The peptide of claim 18 th binding affinity of which is at least thr e-fold greater than that exhibited by the peptide RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

- 91 -

SUBSTITUTE SHEET (RULE 26)

26.   Th   peptid   of claim 1 the binding affinity of
which is at l ast four-fold greater than that exhibited by
the p ptide RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

27.   The peptide of claim 18 the binding affinity of
5 which is at least four-fold greater than that exhibited by
the peptide RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

28.   The peptide of claim 1 which further exhibits a
general binding affinity for the SH3 domain of Abl, Grb2,
PLC-δ, PLC-γ, Ras GAP, Nck, p85 PI-3'Kinase, and proteins
10 related thereto.

29.   The peptide of claim 1 which exhibits a selectiv
binding affinity for the SH3 domain of Src and Src-related
proteins.

30.   The peptide of claim 18 which further exhibits a
15 general binding affinity for the SH3 domain of Abl, Grb2,
PLC-δ, PLC-γ, Ras GAP, Nck, p85 PI-3'Kinase, and proteins
related thereto.

31.   The peptide of claim 18 which exhibits a selective
binding affinity for the SH3 domain of Src and Src-related
20 proteins.

32.   A peptide having the amino acid sequence
RSTPRPLPMLPTTR (SEQ ID NO:62).

33.   A peptide having the amino acid sequence
RSTPRPLPPLPTTR (SEQ ID NO:67).

25   34.   A peptide having the amino acid sequence
GILAPPVPPRNTR (SEQ ID NO:63).

35.   A peptide having the amino acid sequence
VLKRPLPIPPVTR (SEQ ID NO:64).

36.   A peptide having the amino acid sequence
30 GPHRRLPPTPATR (SEQ ID NO:65).

37.   A peptide having the amino acid sequence
ANPSPATRPLPTR (SEQ ID NO:66).

38.   A peptide having an amino acid sequence selected
from the group consisting of RSTRPLPILPRTT, STPRPLPMLPTTR,
35 STNRPLPMIPTTR, RSTRPLPSLPITT, STSRPLPSLPTTR, RSTRSLPPLPPTT,
RSTRQLPIPPTTT, STPRPLPLIPTTP, RSTRPLPPTPLTT, and
RSTRPQPPPPITT (SEQ ID NOS:85-94).

- 92 -

SUBSTITUTE SHEET (RULE 26)

39.   A peptide having the amino acid s quence selected
from th group consisting of VLKRPLPIPPVTR (SEQ ID NO:64),
YSTRPVPPITRPS (SEQ ID NO:76), SHKSRLPPLPTRP (SEQ ID NO:77),
GPHRRLPPTPATR (SEQ ID NO:65), PATRPLPTRPSRT (SEQ ID NO:81),
5 and SGGILAPPVPPRN (SEQ ID NO:84).

40.   A peptide having the amino acid sequence selected
from the group consisting of PPNSPLPPLPTHL (SEQ ID NO:72),
TGRGPLPPLPNDS (SEQ ID NO:74), YRFRALPSPPSAS, LAQRQLPPTPGRD
ALQRRLPRTPPPA (SEQ ID NOS:78-80), YSTRPLPSRPSRT, and
10 XPGRILLLPSEPR (SEQ ID NOS:82-83).

41.   A construct comprising a nucleic acid encoding a
peptide of claim 1 or its complement.

42.   The construct of claim 41 which is a DNA
polynucleotide.

15   43.   The construct of claim 41 which is a RNA
polynucleotide.

44.   A construct comprising a nucleic acid encoding a
peptide of claim 18 or its complement.

45.   The construct of claim 44 which is a DNA
20 polynucleotide.

46.   The construct of claim 44 which is a RNA
polynucleotide.

47.   The construct of claim 41 which is a transforming
vector.

25   48.   The construct of claim 44 which is a transforming
vector.

49.   A host cell transformed with the vector of claim
47.

50.   A host cell transformed with the vector of claim
30 48.

51.   A conjugate comprising a peptide of claim 1 and a
second molecule.

52.   The conjugate of claim 51 in which said second
molecule is selected from the group consisting of an amino
35 acid, a p ptide, a protein, a nucleic acid, a nucleoside, a
glycosidic residue, a label, a drug or a small molecule.

- 93 -

SUBSTITUTE SHEET (RULE 26)

53. A diagnostic kit for the detection of SH3 domains comprising an SH3 domain-binding peptide and a detectable label conjugat d t said peptide directly, indirectly or by complexation, said peptide comprising: (i) a core sequence

5 motif of the formula RXLPφφP (SEQ ID NO:71), wherein X represents any amino acid except cysteine and φ represents a hydrophobic amino acid residue, each letter representing the standard one-letter designation for the corresponding amino acid residue; and (ii) two or more additional amino acid

10 residues flanking said core sequence at its C-terminal end, N-terminal end or both.

54. A drug delivery system comprising an SH3 domain-binding peptide and a drug conjugated to said peptide directly, indirectly or by complexation, said peptide

15 comprising: (i) a core sequence motif of the formula RXLPφφP (SEQ ID NO:71), wherein X represents any amino acid except cysteine and φ represents a hydrophobic amino acid residue, each letter representing the standard one-letter designation for the corresponding amino acid residue; and (ii) two or

20 more additional amino acid residues flanking said core sequence at its C-terminal end, N-terminal end or both.

55. The drug delivery system of claim 54 which may be administered parenterally, orally, enterally, topically or by inhalation.

25 56. The drug delivery system of claim 54 which may b administered intranasally, opthalmically or intravaginally.

57. The drug delivery system of claim 54 which is in the form of a solid, gel, liquid or aerosol.

58. A method of modulating the activity of Src or Src-

30 related proteins comprising administering a composition comprising an effective amount of a peptide of claim 1 and a carrier.

59. The method of claim 58 which inhibits the activity of Src or Src-related proteins.

35 60. The meth d of claim 58 which activates Src or Src-related proteins.

- 94 -

SUBSTITUTE SHEET (RULE 26)

61. A method of identifying a peptide having a region
that binds to an SH3 domain comprising:

(a) providing an immobilized target protein
comprising an SH3 domain;

5   (b) incubating said immobilized target protein
with an aliquot taken from a random peptide library;

(c) washing unbound peptide from said immobilized
target protein;

(d) recovering the peptide bound to said
10 immobilized target protein; and

(e) determining the primary sequence of the SH3
domain-binding peptide.

62. The method of claim 61 in which said library is a
displayed random peptide library.

15   63. The method of claim 62 in which said library is a
phage-displayed random peptide library.

64. The method of claim 62 in which said library is a
phagemid-displayed random peptide library.

65. The method of claim 61 in which step (c) includes
20 washing unbound phage from said immobilized target protein;
step (d) includes recovering the phage bound to said
immobilized target protein; and step (e) includes determining
the relevant nucleotide sequence of said binding phage
nucleic acid, from which the primary sequence corresponding
25 to the SH3 domain-binding peptide is deduced.

66. A method of identifying a peptide having a region
that binds to an SH3 domain comprising:

(a) providing an immobilized target protein
comprising an SH3 domain;

30   (b) incubating said immobilized target protein
with an aliquot taken from a phage-displayed random peptide
library, which library includes peptides having a random
sequence of ≥8 amino acid residues;

(c) washing unbound phage from said immobilized
35 target protein;

(d) recovering th phage bound to said immobilized
target pr tein; and

- 95 -

SUBSTITUTE SHEET (RULE 26)

( ) determining th relevant nucleotide sequenc of said binding phage nucleic acid and deducing the primary sequence corr sponding to the SH3 domain-binding peptid .

67. The method of claim 66 which further comprises
5 amplifying the titer of the recovered phage.

68. The method of claim 66 which further comprises repeating the steps of incubation, washing and recovery to provide SH3 domain-binding peptide-enriched phage.

69. A pharmaceutical composition comprising an SH3
10 domain-binding peptide and a pharmaceutically acceptable carrier, said peptide comprising: (i) a 9-mer sequence motif of the formula RXLPφφPXψ (SEQ ID NO:10), wherein X represents any amino acid except cysteine, φ represents a hydrophobic amino acid residue, and wherein ψ is a hydrophilic amino acid
15 residue except cysteine, each letter representing the standard one-letter designation for the corresponding amino acid residue; and, optionally, (ii) additional amino acid residues flanking said 9-mer sequence at its C-terminal end, N-terminal end or both, up to a total of 45 amino acid
20 residues, including said 9-mer sequence.

70. The composition of claim 69 in which at least one additional amino acid flanks said 9-mer sequence.

71. The composition of claim 69 in which at least two additional amino acids flank said 9-mer sequence.

25 72. The composition of claim 69 in which at least three additional amino acids flank said 9-mer sequence.

73. A method of disrupting protein tyrosine kinase-mediated signal transduction pathways comprising administering an effective amount of a peptide of claim 1.

30 74. A method of regulating the processing, trafficking or translation of RNA by administering an effective amount of a peptide of claim 1.

35

- 96 -

SUBSTITUTE SHEET (RULE 26)

                              Pro.Gly
GGC.TCG.AGN.(NNB)$_{18}$.CCA.GGT

                    GGT.CCA.(NNV)$_{18}$.AGA.TCT.GG

N=A,C,G,T
B=C,G,T   FILL IN WITH Taq DNA POLYMERASE
V=A,C,G

  Xho I
GGC.TCG.AGN.(NNB)$_{18}$.CCA.GGT

                    GGT.CCA.(NNV)$_{18}$.AGA.TCT.GG
                                      Xba I
                    CLEAVE WITH Xho I AND Xba I

    TCG.AGN.(NNB)$_{18}$.CCA.GGT

                    GGT.CCA.(NNV)$_{18}$.AGA.TC

                    LIGATE WITH Xho I + Xba I-
                    CLEAVED M13 m663 VECTOR

                    ELECTROPORATE INTO E. coli DH5α F'

                    TSAR-9 LIBRARY
        OF pIII-RANDOM SEQUENCE FUSION PROTEINS

    . . . S H S | S (R/S) X$_{18}$ P G X$_{18}$ S R | P S R T . . .

SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.1

Gly

tt.ttg.tcg.acN.(NNB)$_{10}$.Ngc.ggt.g

cg.cca.cNV.(NNV)$_{10}$.tga.tca.ttt.t

N=A,G,T,C
B=G,T,C
V=G,A,C

↓ FILL IN WITH Taq DNA POLYMERASE

Sal I

tt.ttg.tcg.acN.(NNB)$_{10}$.Ngc.ggt.g

cg.cca.cNV.(NNV)$_{10}$.tga.tca.ttt.t
Spe I

↓ CLEAVE WITH Sal I + Spe I

tcg.acN.(NNB)$_{10}$.Ngc.ggt.g

cg.cca.cNV.(NNV)$_{10}$.tga.tc

↓ LIGATE WITH Xho I + Xba I-
CLEAVED M13 m663 VECTOR

↓ ELECTROPORATE INTO E. coli DH5α F'

TSAR-12 LIBRARY
OF pIII-RANDOM SEQUENCE FUSION PROTEINS

. . . S H S $\boxed{S \ (S/T) \ X_{10} \phi G \partial X_{10} \ T \ R}$ P S R T . . .

↑

ϕ=S,R,G,C, OR W
∂=V,A,D,E, OR G

SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.2

<u>Xho I</u>

T GAC GTC TCG AGT TGT (NNK)$_8$ TGT GGA TCT AGA AGG ATC

CCT <u>AGA TCT</u> TCC TAG
<u>Xba I</u>

N=A,C,G, OR T
K=G OR T
M=A OR C

FILL IN WITH DNA POLYMERASE

Xho I

Xba I

CLEAVE WITH Xho I AND Xba I

TCG AGT TGT (NNK)$_8$ TGT GGA T

CA ACA (NNM)$_8$ ACA CCT AGA TC

LIGATE INTO m663 VECTOR TREATED WITH Xho I, Xba I, AND CALF INTESTINE ALKALINE PHOSPHATASE

ELECTROPORATE INTO E. coli XL1-B CELLS

S S C (X)$_8$ C G S R P S R T . .

SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.3

Xba I

TCG AGT TGT (NNK)$_8$ TGT GGA T CT AGA TCC ACA (MNN)$_8$ ACA AC

CA ACA (NNM)$_8$ ACA CCT A GA TCT AGG TGT (KNN)$_8$ TGT TGA GCT

Xho I   Xho I

LIGATION OF TWO DOUBLE-STRANDED
OLIGONUCLEOTIDES INTO A HEAD-TO-HEAD
ARRANGEMENT AT THE Xba I SITE

INSERTION INTO m663 VECTOR CLEAVED
BY Xho I AND Xba I. ILLEGITIMATE LIGATION
AT THE Xba I SITE OF THE VECTOR.

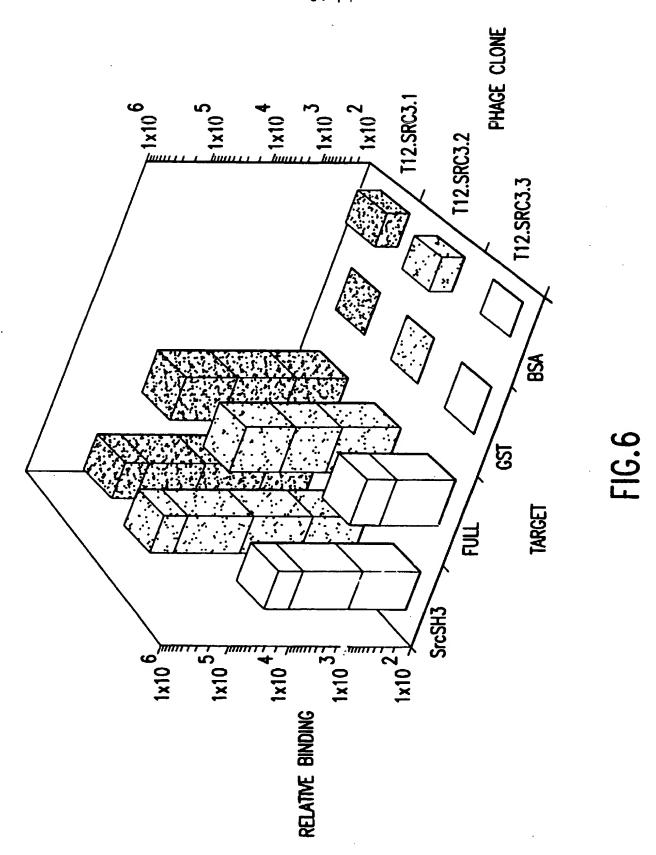ELECTROPORATE INTO E. coli XLI-B CELLS

S S C (X)$_8$ C G S R S T (X)$_8$ T T R . .

SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.4

FIG.5

CLONE	SEQUENCE			FREQUENCY
T9.SRC3.2	SSFDQQDWDYSIAEKMHPIRPGF	RELPPLP	PSRASFGGGASRPSR	2
T12.SRC3.4	STNVVVTGSVIARGAQS	RPLPIPP	EIRPSR	1
T12.SRC3.6	STAPVGLRVAHEGGVLK	RPLPIPP	VTRPSR	1
T9.SRC3.4	SSSGYVVPKRLGDMREYNAHPGLHVPPN	SPLPPLP	THLQSSRPSR	1
T9.SRC3.6	SSRGEGNNIISSRPFLSNSDPGVSNKLTGR	GPLPPLP	NDSRPSR	2
T12.SRC3.7	STAVSFRFMPGGGGAFYST	RPVPPIL	RPSRT	1
T12.SRC3.5	STAHSLVDWGTFSGVSHKS	RLPPLPT	RPSRT	1
T9.SRC3.7	XPGYARIVSYRF	RALPSPP	SASRPSR	1
T12.SRC3.3	SINDVDVWHHWWNSGGPH	RRLPPTP	ATRPSR	4
T9.SRC3.5	SSDNVARRVHASELIYTDLSPGILLAQ	RQLPPTP	GRDPSHSRPSR	1
T9.SRC3.1	SSESPLMYNRVGALQSLTSVPGSMHHFALQ	RRLPRTP	PPASRPSR	5
T12.SRC3.2	STRVSHSVPGYVGGANPSPAT	RPLPTRP	SRTVESC	19
T9.SRC3.3	SRYNDLGTRPVSC"IKYDYFPGYSQHVITPDGSYST	RPLPSRP	SRTVESC	2
T9.SRC3.8	XPG	RILLLPS	EPRTFYNYGHDSRPSR	1
T12.SRC3.1	SIMYGVSWLSSGSGGILA	PPVPPRN	TRPSR	25
Consensus		RPLPPLP		

CLONE	SEQUENCE			FREQUENCY
R8C.YES3.6	SSCTEKTVSGWCGSRST	RPLPILP	RTTRPSR	1
R8C.YES3.5	SSCMLPTDGWQCGSRSTP	RPLPMLP	TTRPSR	1
R8C.YES3.3/SRC3.1	SSCDGTQFRLNCGSRSTN	RPLPMIP	TTRPSR	3
R8C.YES3.1/SRC3.2	SSCMGQQAGLKCGSRST	RPLPSLP	ITTRPSR	7
R8C.YES3.7	SSCYREKDTWGCGSRSTS	RPLPSLP	TTRPSR	2
R8C.YES3.2	SSCLFEQGAGTCGSRST	RSLPPLP	PTTRPSR	2
R8C.YES3.4	SSCDHTLGLGWCGSRST	RQLPIPP	TTIRPSR	1
R8C.YES3.10	SSCDTGRIAPGCGSRSTP	RPLPLIP	TTPRSTNLNLTSTTTRPSR	2
R8C.YES3.8	SSCGLDNAAKTCGSRST	RPLPPTP	LTTRPSR	2
R8C.YES3.9	SSCSRAHETEMCGSRST	RPQPPPP	ITTRPSR	2
Consensus		RPLPPLP		

FIG.6

OD ELISA VALUES

FIG.7

OD ELISA VALUES

FIG.8

FIG.9

PEPTIDES:
(MICROMOLAR
CONCENTRATION)

	GST		
	Src		
	1	10	100

12.SRC3.1   GILAPPVPPRNTR

12.SRC3.6   VLKRPLPIPPVTR

LONG CONSENSUS RST P RPLPPLPTR

PI3 KINASE   KISPPTPKPRPPRPLPV

VINCULIN   LAPPKPPLPEGEV

	Abl	
	100	

12.SRC3.1   GILAPPVPPRNTR
12.SRC3.6   VLKRPLPIPPVTR
LONG CONSENSUS   RSTPPRRPLPPLPTTR
PI3 KINASE   KISPPTPKPRPPRPLPV

FIG.10

FIG.11A

FIG.11B

FIG.11C

FIG.11D

ANNEAL OLIGONUCLEOTIDES

FILL IN

DIGEST WITH
Xho I AND Xba I

GEL PURIFY
CUT INSERT

CLONE INTO
nBAX VECTOR

EXPRESS BPL

CTG TGC CTC GAG K NNK$_6$ CCA NNK$_2$ TCT AGA CGT GTC AGT
　　　　　　　　　　　　　　　　　　　AGA TCT GCA CAG TCA

CTG TGC CTC GAG K NNK$_6$ CCA NNK$_2$ TCT AGA CGT GTC AGT
GAC ACG GAG CTC M NNM$_6$ GGT NNM$_2$ AGA TCT GCA CAG TCA

CTG TGC C　TC GAG K NNK$_6$ CCA NNK$_2$ CCA NNK$_6$ T　　CT AGA CGT GTC AGT
GAC GAG CT　C M NNM$_6$ GGT NNM$_2$ AGA TC　T GCA CAG TCA

TC GAG K NNK$_6$ CCA NNK$_2$ CCA NNK$_6$ T
　　 C M NNM$_6$ GGT NNM$_2$ GGT NNM$_6$ AGA TC

. . .TC GAG K NNK$_6$ CCA NNK$_2$ CCA NNK$_6$ TCT AGA. . . .
. . .AG CTC M NNM G　GGT NNM$_2$ GGT NNM$_6$ AGA CTC. . . . .

　　　　　S S/R　X$_6$P　X$_2$　P　X$_6$　S　R . . . . pIII

PEPTIDE CLEAVAGE
SITE

FIG.12

FIG.13

mBAX

SIGNAL PEPTIDE
CLEAVAGE SITE

-1 ↓ +1

S S │ I D M P * T A S T M Y N M L H R N E P │
. . TCCTCGAGTATCGACATGCCTTAGACTGCTAGCACTATGTACAACATGCTTCATGCAACGAGCCA . .
        Xho I

. . . epitope,mAb . . .

│ G G R K L S P P A N D M P P A L L K R │ S R
. . GGTGGGAGGAAGTTGAGCCCGCCGCCAACGACATGCCGCCCTCCTGAAGAGGTCTAGA . .
                                              Xba I

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.

US CL :435/5, 7.1, 172.1, 320.1; 530/300; 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/5, 7.1, 172.1, 320.1; 530/300; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG, STN

search terms: sh3, src, peptides

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	The Journal of Biological Chemistry, Volume 269, Number 39, issued 30 Spetember 1994, Sparks et al., "Identification and Chracterization of Src SH3 Ligands from Phage-displayed Random Peptide Libraries", pages 23853-23856, see entire article.	1-74
X --- Y	Journal of the American Chemical Society, Volume 115, issued 1993, Chen et al. "Biased Combinatorial Libraries: Novel Ligands for the SH3 Domain of Phosphatidyllinositol 3-Kinase", pages 12591-12592, see entire article, especially Tables I and II.	1-6, 24, 26, 28, 29, 53 ------------ 7-23, 25, 27, 30-51, 53-74

[X] Further documents are listed in the continuation of Box C.   [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 OCTOBER 1995	01 NOV 1995

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	LORA M. GREEN
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Science, Volume 249, issued 28 September 1990, Langer, "New Methods of Drug Discovery", pages 1527-1533, see entire article.	51, 52, 54-57
Y	Science, Volume 257, issued 07 August 1992, Cicchetti et al., "Identificat ion of a Protein That Binds to the SH3 Region of Abl and Is Similar to BCR and GAP-rho"; pages 803-806, see entire article.	7-23, 25, 27, 30-50
Y	Gene, Volume 128, issued 1993, Kay et al., "An M13 Phage Library Displaying Random 38-amino-acid Petides as a Source of Novel Sequences with Affinity to Selected Targets", pages 59-65, see entire article.	7-23, 25, 27, 30-48, 61-68

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

G01N 33/53; C12Q 1/70; C12N 5/10. 7/01; C07K 7/00, 14/00; A61K 38/03

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 33/53, C12Q 1/70, C12N 5/10, 7/01, C07K 7/00, 14/00, A61K 38/03	A1	(11) International Publication Number: WO 96/03649
		(43) International Publication Date: 8 February 1996 (08.02.96)

(21) International Application Number: PCT/US95/09382

(22) International Filing Date: 24 July 1995 (24.07.95)

(30) Priority Data:
278,865   22 July 1994 (22.07.94)   US
483,555   7 June 1995 (07.06.95)   US

(71) Applicant: THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; Office of Technology Development, CB #4100, 302 Bynum Hall, Chapel Hill, NC 27599-4100 (US).

(72) Inventors: SPARKS, Andrew, B.; 201 Blue Ridge, Carrboro, NC 27510 (US). KAY, Brian, K.; 18 Wysteria Way, Chapel Hill, NC 27514 (US). THORN, Judith, M.; Apartment 18-G, 200 Barnes, Carrboro, NC 27510 (US). QUILLIAM, Lawrence, A.; Apartment 5, 206 Pinegate Circle, Chapel Hill, NC 27514 (US). DER, Channing, J.; 101 Fieldstone Court, Chapel Hill, NC 27514 (US).

(74) Agents: MISROCK, S., Leslie et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.

(54) Title: Src SH3 BINDING PEPTIDES AND METHODS OF ISOLATING AND USING SAME

(57) Abstract

Peptides having general and specific binding affinities for the Src homology region 3 (SH3) domains of proteins are disclosed in the present invention. In particular, SH3 binding peptides have been isolated from three phage-displayed random peptide libraries which had been screened for isolates that bind to bacterial fusion proteins comprising SH3 and glutathione S-transferase (GST). Preferred peptides are disclosed which comprise a core 7-mer sequence (preferably, a consensus motif) and two or more, preferably at least six, additional amino acid residues flanking the core sequence, for a total length of 9, preferably at least 13, amino acid residues and no more than about 45 amino acid residues. Such peptides manifest preferential binding affinities for certain SH3 domains. The preferred peptides exhibit specific binding affinities for the Src-family of proteins. *In vitro* and *in vivo* results are presented which demonstrate the biochemical activity of such peptides.

Src SH3 BINDING PEPTIDES AND
METHODS OF ISOLATING AND USING SAME

1.   Field of the Invention

5   The present invention relates to SH3 binding peptides having a broad range of binding specificities. That is, certain members of the SH3 binding peptides disclosed bind with approximately the same facility with SH3 domains derived from different SH3 domain-containing proteins. Oth r

10 members, in contrast, bind with a much greater degree of affinity for specific SH3 domains. The SH3 binding peptides are obtained from random peptide libraries that are also phage-displayed. Methods are described of obtaining the phage clones that bind to the SH3 domain targets and of

15 determining their relevant nucleotide sequences and consequent primary amino acid sequence of the binding peptides. The resulting SH3 binding proteins are useful in a number of ways, including, but not limited to, providing a method of modulating signal transduction pathways at th

20 cellular level, of modulating oncogenic protein activity or of providing lead compounds for development of drugs with th ability to modulate broad classes, as well as specific classes, of proteins involved in signal transduction.

25   ### 2.   Background of the Invention
        #### 2.1. Src and the SH3 Domain
                Among a number of proteins involved in eukaryotic cell signaling, there is a common sequence m tif called the SH3 domain. It is 50-70 amino acids in length,

30 moderately conserved in primary structure, and can be pr s nt from one to several times in a large number of proteins involved in signal transduction and in cytoskeletal proteins.
        The protein pp60c-src represents a family of at least nine non-receptor protein tyrosine kinases (NR-PTKs).

35 Members f this family share an verall structural organization comprising a series f catalytic and n n-catalytic d mains. In Src, a 14-amin -acid myristylation

- 1 -

signal r sid s at th   xtreme amino-terminus, and is follow d
by a uniqu r gion that is n t highly cons rved among family
members. Following this region ar two highly c nserved 60-
and 100-amino-acid regions, the Src homology (SH) domains 3
5 and 2, respectively. SH2 and SH3 domains have been shown to
play an important role in mediating protein-protein
interactions in a variety of signaling pathways. Koch, C.A.,
et al., in Science (1991) 252:668-74. The carboxy-terminal
half of Src contains the PTK catalytic domain, as well as a
10 negative regulatory tyrosine (Y527) near the carboxy
terminus. Phosphorylation of this residue (e.g., by Csk)
results in the inhibition of PTK activity. Cooper, J.A., et
al., in Science (1986) 231:1431-1434. Mutation of Y527->F
generates forms of Src with increased PTK and oncogenic
15 activity. Cartwright, C.A., et al., in Cell (1987) 49:83-91;
Kmiecik, T.E., et al., in Cell (1987) 49:65-73; and Piwicna-
Worms, H., et al., in Cell (1987) 75-82.

      The fact that some mutations which result in increased
Src PTK and transforming activity map to the Src SH2 (Seidel-
20 Dugan, C., et al., in Mol. Cell. Biol. (1992) 12:1835-45; and
Hirai, H. and Varmus, H.E. in Mol. Cell. Biol. (1990)
10:1307-1318) and SH3 domains (Seidel-Dugan, C., et al.,
supra; Hirai, H. and Varmus, H.E., supra; Superti-Furga, G.,
et al., in Embo. J. (1993) 12:2625-34; and Potts, W.M., et
25 al., in Oncogene Res. (1988) 3:343-355) suggests a negative
regulatory role for these domains. That phosphotyrosine
residues within specific sequence contexts represent high
affinity ligands for SH2 domains suggests a model in which
the SH2 domain participates in Y527-mediated inhibition of
30 PTK activity by binding phosphorylated Y527, thereby locking
the kinase domain in an inactive configuration. Matsuda, M.,
Mayer, B.J., et al., in Science (1990) 248:1537-1539. This
model is supported by the observation that phosphopeptides
corresponding to the carboxy-terminal tail of Src bind
35 active, but not inactiv , variants of Src. R uss 1, R.R., et
al., in Proc. Natl. Acad. Sci. U S A (1991) 88:10696-700; and
Liu, X., et al., in Oncogene (1993) 8:1119-1126.

- 2 -

Th m chanism f SH3-m diated inhibition of Src PTK
activity remains unclear. There is evid nc that pY527-
mediat d inhibiti n of Src PTK activity involv s th SH3
domain as well as the SH2 domain. Okada, M., Howell, et al.,
5 in _J. Biol. Chem._ (1993) 268:18070-5; Murphy, S.M., et al.,
in _Mol. Cell. Biol._ (1993) 13:5290-300; and Superti-Furga,
G., et al., _supra_. Although these effects are thought to be
a consequence of SH3-mediated protein-protein interactions,
precisely how the Src SH3 domain exerts its negative
10 regulatory effect is unclear. Identification of high
affinity ligands for the Src SH3 domain could help resolve
these issues.

2.2. Protein Tyrosine Kinases and The Immune Response

15   Src-related tyrosine kinases are expressed in a
variety of cell types including those of the immune system
(lymphocytes, T cells, B cells, and natural killer cells) and
the central nervous system (neural cells, neurons,
oligodendrocytes, parts of the cerebellum, and the like).
20 Umemori, H. et al., in _Brain Res. Mol. Brain Res._ (1992) Dec.
16(3-4):303-310. Their presence in these cells and tissu s
and their interaction with specific cell surface receptors
and immunomodulatory proteins (such as T cell antigen
receptor, CD14, CD2, CD4, CD40 or CD45) suggest that these
25 kinases serve an important role in the signalling pathways of
not only the central nervous system but of the immune system,
as well. See, e.g., Ren, C.L. et al., in _J. Exp. Med._ (1994)
179(2):673-680 (signal transduction via CD40 involves
activation of Lyn kinase); Donovan, J.A. and Koretzky, G.A.,
30 in _J. Am. Soc. Nephrol._ (1993) 4(4):976-985 (CD45, the immune
response, and regulation of Lck and Fyn kinases); and Carmo,
A.M. et al., in _Eur. J. Immunol._ (1993) 23(9):2196-2201
(physical association of the cytoplasmic domain of CD2 with
p56lck and p59fyn).
35   F r instanc , mic with disruptions in their Src-like
genes, Hck and Fgr, possess macrophag s with impair d
phagocytic activity or xhibit a novel immunodeficiency

- 3 -

SUBSTITUTE SHEET (RULE 26)

characteriz d by an incr as d susceptibility to infecti n
with List ria m nocytog n s. Lowell, C.A. t al., in <u>Genes</u>
<u>Dev.</u> (1994) 8(4):387-398. Also, it has been sh_wn that
bacterial lipopolysaccharide (LPS) activates CD14-associated
5 p56lyn, p68hck, and p59c-fgr, while inducing the production
of lymphokines, such as TNF-alpha, IL-1, IL-6, and IL-8.
Inhibition of the protein tyrosine kinases blocks production
of TNF-alpha and IL-1.

10   **2.3. SH3 Binding Peptides**

         As mentioned above, it has long been suspected that
SH3 domains are sites of protein-protein interaction, but it
has been unclear what SH3 domains actually bind. Efforts to
identify ligands for SH3 domains have led to the
15 characterization of a number of SH3-binding proteins,
including 3BP1 and 2 (Ren, R., Mayer, et al., in <u>Science</u>
(1993) 259:1157-61), SOS (Olivier, J.P., et al., in <u>Cell</u>
(1993) 73:179-91; and Rozakis-Adcock, M., et al., in <u>Nature</u>
(1993) 363:83-5), p85 PI-3' Kinase (Xingquan, L., et al., in
20 <u>Mol. Cell. Biol.</u> (1993) 13:5225-5232), dynamin (Gout, I., t
al., in <u>Cell</u> (1993) 75:25-36), AFAP-110 (Flynn, D.C., et al.,
in <u>Mol. Cell. Biol.</u> (1993) 13:7892-7900), and CD42 (Barfod,
E.T., et al., in <u>J. Biol. Chem.</u> (1993) 268:26059-26062).
These proteins tend to possess short, proline-rich stretches
25 of amino acids, some of which have been directly implicated
in SH3 binding. A variety of consensus sequences have been
proposed, although the similarity among proline-rich regions
of different SH3-binding proteins tends to be fairly low.
Also, attempts to build consensus sequences are likely
30 complicated by the incorporation of data from proteins that
bind different SH3 domains.
        Thus, Cicchetti, P., et al., in <u>Science</u> (1992) 257:803-
806, published their work relating to the isolation and
sequencing of two naturally-occurring proteins that could be
35 bound *in vitro* by th SH3 domain f the *abl* oncog n product.
These work rs found that SH3 domains bind short, proline-rich
regions of such proteins. Subsequ ntly, this same group

- 4 -

disclosed furth r r sults (Ren, R. et al., *supra*) in which
th SH3 binding sit s f the SH3 binding prot ins w r
localized to "a nine- or ten-amino acid stretch rich in
proline residues." A consensus sequence incorporating the
5 features of the SH3 binding sites of four SH3 binding
proteins was proposed: XPXXPPP♥XP (SEQ ID NO:1), wherein X
indicates a position in the amino acid sequence which is n t
conserved among the four SH3 binding proteins, P represents
proline, and ♥ indicates a hydrophobic amino acid residue,
10 such as P or L.

The screening of complex random peptide libraries has
been used to identify peptide epitopes for monoclonal (Scott,
J.K. and Smith, G.P. in Science (1990) 249:386-390) and
polyclonal (Kay, B.K., et al., in Gene (1993) 128:59-65)
15 antibodies, as well as peptide ligands for a variety of
proteins, including streptavidin (Devlin, J.J., et al., in
Science (1990) 249:404-406; and Lam, K., et al., in Nature
(1991) 354:82-84), the endoplasmic reticulum chaperone BiP
(Blond-Elguindi, S., et al., in Cell (1993) 75:717-728), and
20 CaM (Dedman, J.R., et al., in J. Biol. Chem. (1993)
268:23025-23030).

Recently, Chen, J.K. et al., in J. Am. Chem. Soc. (1993)
115:12591-12592, described ligands for the SH3 domain of
phosphatidylinositol 3-kinase (PI-3' Kinase) which were
25 isolated from a biased combinatorial library. A "biased"
library is to be distinguished from a "random" library in
that the amino acid residue at certain positions of the
synthetic peptide are fixed, i.e., not allowed to vary in a
random fashion. Indeed, as stated by these research workers,
30 screening of a "random" combinatorial library failed to yield
suitable ligands for a PI-3' Kinase SH3 domain probe. The
binding affinities of these unsuitable ligands was described
as weak, >100 μM, based on dissociation constants measured by
the Biosensor System (BIAcore).

35

Mor recently, Yu, et al. (Yu, H., et al., in Cell
(1994) 76:933-945) used a "bias d" synthetic peptide library

- 5 -

SUBSTITUTE SHEET (RULE 26)

f th form XXXPPXPXX (SEQ ID NO:2), wherein X represents any
amino acid ther than cysteine, to id ntify a series of
peptid s which bind the Src and PI-3' Kinase SH3 domains.
The bias was accomplished by fixing the proline residues at
5 the specific amino acid positions indicated for the "random"
peptide. As stated previously, without this bias, the
technique disclosed fails to identify any SH3 domain-binding
peptides.

A consensus sequence, based on 13 binding peptides was
10 suggested: RXLPPRPXX (SEQ ID NO:3), where X tends to be a
basic residue (like R, K or H). The binding affinities of
several SH3 binding peptides were disclosed as ranging from
8.7 to 30 μM. A "composite" peptide, RKLPPRPRR (SEQ ID
NO:4), was reported to have a binding affinity of 7.6 μM.
15 This value compares favorably to the binding affinity of the
peptide, VPPPVPPRRR (SEQ ID NO:5), to the N-terminal SH3
domain of Grb2. See, Kraulis, P.J. _J. Appl. Crystallogr_.
(1991) 24:946. Recognizing the limitations of their
technique, Chen and co-workers, _supra_, stated that their
20 results "illustrate the utility of biased combinatorial
libraries for ligand discovery in systems _where there is some
general knowledge of the ligand-binding characteristics of
the receptor_" (emphasis added).

Yu and co-workers, _supra_, further described an SH3
25 binding site consensus sequence, XpØPpXP (SEQ ID NO:6),
wherein X represents non-conserved residues, Ø represents
hydrophobic residues, P is proline, and p represents residues
that tend to be proline. A consensus motif of RXLPPRPXX (SEQ
ID NO:7), where X represents any amino acid other than
30 cysteine, was proposed for ligands of PI-3' Kinase SH3
domain. A consensus motif of RXLPPLPRϕ (SEQ ID NO:8), where
ϕ represents _hydrophobic_ residues, was proposed for ligands
of Src SH3 domain. Still, the dissociation constants
reported for the 9-mer peptides ranged only from about 8-70
35 μM and s l ctivity between one type of SH3 domain and another
was relatively poor, the K_Ds diff ring by only about a factor
of four.

- 6 -

SUBSTITUTE SHEET (RULE 26)

Henc , there remains a n ed to d v lop techniques for
th id ntification of Src SH3 binding peptid s which do not
rely on such "bias d" combinatorial peptide libraries that
are limited to a partially predetermined set of amino acid
5   sequences. Indeed, the isolation of SH3 binding peptides
from a "random" peptide library has not been achieved
successfully before now. Furthermore, particular peptides
having much greater binding affinities, whether general or
more selective binding for specific SH3 domains, remain to be
10 identified. Binding peptides specific for particular SH3
domains are useful, for example, in modulating the activity
of a particular SH3 domain-containing protein, while leaving
others bearing an SH3 domain unaffected. Still, the more
promiscuous general binding peptides are useful for the
15 modulation of a broad spectrum of SH3 domain-containing
proteins.

The present invention relates to such SH3 binding
peptides, methods for their identification, and compositions
comprising same. In particular, peptides comprising
20 particular sequences of amino acid residues are disclosed
which were isolated from random peptide libraries. In the
present invention, clones were isolated from a phage-
displayed random peptide library which exhibited strong
binding affinities for SH3 domain-containing protein targets.
25 Some of these protein targets, include Abl, Src, Grb2, PLC-δ,
PLC-γ, Ras GAP, Nck, and p85 PI-3' Kinase. From the
nucleotide sequence of the binding phage, the amino acid
sequence of the peptide inserts has been deduced. Synthetic
peptides having the desired amino acid sequences are shown to
30 bind the SH3 domain of the target proteins. In particular,
synthetic peptides combining a core consensus sequence and
additional amino acid residues flanking the core sequence are
especially effective at binding to particular target protein
SH3 domains. The SH3 binding peptides disclosed herein can
35 be utilized in a number f ways, including the p t ntial
modulation f oncogenic prot in activity *in vivo*. These
peptid s also serv as useful l ads in the production of

- 7 -

SUBSTITUTE SHEET (RULE 26)

p ptidomim tic drugs that modulat a large class of proteins involved in signal transducti n pathways and oncogenesis.

3. Summary of the Invention

5   Accordingly, three phage-displayed random peptid libraries were screened for isolates that bind to bacterial fusion proteins consisting of the Src homology region 3 (SH3) and glutathione S-transferase (GST). DNA sequencing of the isolates showed that they contained sequences that resemble

10 the consensus motif, RPLPPLP (SEQ ID NO:9), within their 8, 22, or 36 amino acid long random regions. When peptides were synthesized corresponding to the pIII inserts of the SH3-binding phage, they bound to the GST fusions of the SH3 domains of Src and the Src-related proteins, such as Yes, but

15 not of Grb2, Crk, Abl, or PLCγ1. The synthesized peptides bind quite well to the Src SH3 domain and act as potent competitors of natural Src SH3 interactions in cell lysates. For instance, these peptides can compete with radiolabelled proteins from cell lysates in binding to immobilized Src-GST,

20 with an apparent IC_{50} of 1-10 μM. When a peptide, bearing the consensus sequence RPLPPLP (SEQ ID NO:9) was injected into *Xenopus laevis* oocytes, it accelerated the rate of progesterone-induced maturation. These results demonstrate the utility of phage-displayed random peptide libraries in

25 identifying SH3-binding peptide sequences and that such identified peptides exhibit both *in vivo* and *in vitro* biological activity.

        Thus, it is an object of the present invention to provide peptides having at least nine and up to forty-five

30 amino acid residues, including an amino acid sequence of th formula, R-2-L-P-5-6-P-8-9 (SEQ ID NO:10), positioned anywhere along the peptide, in which each number represents an amino acid residue, such that 2 represents any amino acid residue except cysteine, 5 and 6 each represents a

35 hydrophobic amino acid residu , 8 r pres nts any amino acid residue exc pt cystein , and 9 repr sents a hydrophilic amino acid residu except cystein , each l tter being the standard

- 8 -

SUBSTITUTE SHEET (RULE 26)

one-l tter symbol for the c rresponding amino acid, said
p ptide exhibiting a binding affinity for th SH3 d main of
Src, provid d that said peptide is n t R-P-L-P-P-L-P-T-S (SEQ
ID NO:11). In a particular embodiment of the present
5 invention, the peptides also exhibit a binding affinity for
the SH3 domain of Src-related proteins, including Yes, Fyn,
Lyn, Lck, Hck and Fgr.

The present invention also contemplates SH3 domain-
binding peptides that further comprise a C-terminal-flanking
10 amino acid sequence of the formula 10, 10-11, 10-11-12, 10-
11-12-13 (SEQ ID NO:12) or 10-11-12-13-14 (SEQ ID NO:13), in
which each number represents any amino acid residue except
cysteine, such that 10 is bound to 9 by a peptide bond.
Furthermore, peptides are also provided which further
15 comprise an N-terminal-flanking amino acid sequence of the
formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in
which each number represents any amino acid residue except
cysteine, such that 1' is bound to R by a peptide bond.

Thus, in a particular embodiment, a peptide is disclosed
20 having at least thirteen and up to forty-five amino acid
residues, including an amino acid sequence of the formula,
3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned
anywhere along the peptide, in which each number represents
an amino acid residue, such that 3', 2', 1', 2, 8, and 10
25 each represents any amino acid residue except cysteine, 5 and
6 each represents a hydrophobic amino acid residue, and 9
represents a hydrophilic amino acid residue except cysteine,
each letter being the standard one-letter symbol for the
corresponding amino acid, said peptide exhibiting a binding
30 affinity for the SH3 domain of Src.

The present invention also seeks to provide new
consensus sequences or motifs that reflect variations in SH3
domain binding selectivities or specificities. The present
invention also contemplates conjugates of the SH3 binding
35 peptid s and a s cond mol cul or chemical moiety. This
s cond molecule may be any desir d substance whose delivery
to the regi n of th SH3 domain of a particular prot in (or

- 9 -

SUBSTITUTE SHEET (RULE 26)

c ll c ntaining the protein) is s ught. Possible target
c lls include, but are not limited to, n ural cells, immune
cells (e.g., T cells, B cells, natural killer c lls, and th
like), osteoclasts, platelets, epidermal cells, and the like,
5 which cells express Src, Src-related proteins, and
potentially, other SH3 domain-containing proteins. In this
manner, the modulation of the biological activity of proteins
bearing an SH3 domain can be accomplished.

    Other methods and compositions consistent with the
10 objectives of the present invention are likewise disclosed.
In particular, a method is disclosed of modulating the
activity of Src or Src-related proteins comprising
administering a composition comprising an effective amount of
a peptide of the present invention and a carrier, preferably
15 a pharmaceutically acceptable carrier. In a specific
embodiment, the contemplated method results in the inhibition
of the activity of Src or Src-related proteins.
Alternatively, the method is effective to activate Src or
Src-related proteins.

20   In yet another embodiment, a method is disclosed of
identifying a peptide having a region that binds to an SH3
domain comprising: (a) providing an immobilized target
protein comprising an SH3 domain; (b) incubating the
immobilized target protein with an aliquot taken from a
25 random peptide library; (c) washing unbound library peptides
from the immobilized target protein; (d) recovering the
peptide bound to the immobilized target protein; and (e)
determining the primary sequence of the SH3 domain-binding
peptide.

30   Moreover, a method is disclosed of imaging cells,
tissues, and organs in which Src or Src-related proteins are
expressed, which comprises administering an effective amount
of a composition comprising an SH3 domain-binding peptide
conjugated to detectable label or an imaging agent.

35   Other obj ctives of th present invention will become
apparent to one of ordinary skill in th art after

- 10 -

SUBSTITUTE SHEET (RULE 26)

consideration of th abov discl sure and the foll wing
detailed descripti n of th pref rr d embodim nts.

4.   Brief Description of the Figures

5   FIG. 1 illustrates a scheme for the generation of a
random 36 amino acid peptide library (TSAR-9; e.g., SEQ ID
NO:16). Oligonucleotides were synthesized (SEQ ID NOS:17-
18), converted into double-stranded DNA, cleaved with
restriction enzymes (SEQ ID NOS:19-20), and cloned into the
10 M13 vector, m663. The random peptide region encoded by the
oligonucleotides is shown in the box (SEQ ID NO:16) and is
situated at the N-terminus of mature protein III (SEQ ID
NO:21).

         FIG. 2 illustrates a scheme for the generation of a
15 random 22 amino acid peptide library (TSAR-12; e.g., SEQ ID
NO:23). Oligonucleotides were synthesized (SEQ ID NOS:24-
25), converted into double-stranded DNA, cleaved with
restriction enzymes (SEQ ID NOS:26-27), and cloned into the
M13 vector, m663. The random peptide region encoded by the
20 oligonucleotides is shown in the box (SEQ ID NO:23) and is
situated at the N-terminus of mature protein III (SEQ ID
NO:28).

         FIG. 3 illustrates a scheme for the generation of a
random 8 amino acid peptide library (R8C; SEQ ID NO:30).
25 Oligonucleotides were synthesized (SEQ ID NOS:31-32),
converted into double-stranded DNA, cleaved with restriction
enzymes (SEQ ID NOS:33-34), and cloned into the M13 vector,
m663. The random peptide region (SEQ ID NO:30) is flanked by
cysteine residues and is situated at the N-terminus of mature
30 protein III (SEQ ID NO:35).

         FIG. 4 illustrates the possible origin of one class of
double-insert R8C recombinants (e.g., encoding SEQ ID NO:36).
Double-stranded oligonucleotides (e.g., SEQ ID NO:37) may
have ligated in a head-to-head fashion at the *Xba* I site
35 pri r t cloning in the *Xho* I- *Xba* I cl av d M13 v ctor.

         FIG. 5 sh ws a list of random p ptide recombinants (SEQ
ID NOS:38-61 and 106) isolat d by the method of the pres nt

- 11 -

inv nti n and the display d peptid s qu nce. The amin acid
sequences ar aligned to highlight the core sequenc s. The
flanking sequences ar shown to the N-terminal and C-t rminal
ends of the core sequence.

5   FIG. 6 graphically illustrates the relative binding
affinities of selected phage clones for various SH3 domains.
The results indicate that certain amino acid sequences
provide generic SH3 domain binding, while others can provide
greater selectivity for the SH3 domain of Src. Still other
10 clones exhibit Src SH3 domain preferential binding.

          FIG. 7 shows the binding of synthetic peptides (SEQ ID
NOS:9 and 62-70) representing Src SH3-selected phage inserts
to Src SH3-GST fusion target (shaded columns) over background
GST binding (unshaded columns) relative to the core peptide
15 RPLPPLP (SEQ ID NO:9) and proline-rich peptide segments
derived from naturally occurring proteins. Bound
biotinylated peptide was detected with streptavidin-alkaline
phosphatase ELISA. Each point was performed in triplicate;
average absorbance at 405 nm is presented. Error bars
20 represent SD.

          FIG. 8 illustrates the relative specificity of selected
peptides (SEQ ID NOS:9 and 62-70) for SH3 domains derived
from different proteins. In particular, the binding
affinities of the peptides for the SH3 domains of the
25 following protein fusion targets were tested: Src SH3-GST,
Yes SH3-GST, Grb2-GST, Crk SH3-GST, Abl SH3-GST, PLCγ1
SH2SH3-GST. Bound biotinylated peptide was detected with
streptavidin-alkaline phosphatase. Each point was performed
in triplicate; values are average signal (absorbance at 405
30 nm) above GST background, with error bars representing
standard deviation. Hatched bars indicate saturation of the
ELISA signal.

          FIG. 9 presents the results of competition experiments
in which selected peptides were found to inhibit the binding
35 of pr t ins from cell lysates to immobilized Src SH3-GST r
Abl SH3-GST protein fusion targets.

- 12 -

SUBSTITUTE SHEET (RULE 26)

FIG. 10 pr sents a graph illustrating the incr as d rate
f progest rone-induced maturati n of oocytes inj cted with
an SH3 domain-binding p ptid , VLKRPLPIPPVTR (SEQ ID NO:64),
of the present invention. Briefly, Stage VI oocyted were·
5 prepared and injected as previously described (see, Kay,
B.K., in Methods in Cell Biol. (1991) 36:663-669). Oocytes
were injected with 40 nL of 100 μM test peptide or water.
After injection, the oocytes were placed in 2 μg/mL
progesterone (Sigma, St. Louis, MO) and scored hourly for
10 germinal vesicle breakdown (GVBD).

FIG. 11 shows the results of fluorescence experiments in
which certain peptides, Panel A = VLKRPLPIPPVTR (SEQ ID
NO:64), Panel B = GILAPPVPPRNTR (SEQ ID NO:63), Panel C =
RSTPRPLPPLPTTR (SEQ ID NO:67), of the invention were shown to
15 localize within cellular compartments thought to contain Src
or Src-related proteins.

FIG. 12 illustrates a scheme for the generation of a
biased peptide library. Oligonucleotides were synthesized,
converted into double-stranded DNA, cleaved with restriction
20 enzymes XhoI and XbaI, and cloned into the mBAX vector,
described further below in the Examples section. The bias d
peptide region is situated at the N-terminus of mature pIII
protein.

FIG. 13 illustrates the peptide sequence encoded in the
25 mBAX vector situated at the N-terminus of mature pIII
protein.

5.   Detailed Description of the Preferred Embodiments
      5.1. General Considerations
30   The present invention relates to peptides that
exhibit a binding affinity for an SH3 domain, which domain
has been found to be present in a number of physiologically
significant proteins. In particular, peptides are disclosed
which exhibit general binding characteristics to the SH3
35 domains found in a group f prot ins, including but not
limit d t Abl, Src, Grb2, PLC-δ, PLC-γ, Ras GAP, Nck, and
p85 PI-3' Kinase. Pref rr d peptid s exhibit sel ctive, if

- 13 -

SUBSTITUTE SHEET (RULE 26)

not specific, binding affinity for the SH3 domain of Src.   As
d scribed herein, th peptides of the pr s nt inventi n
include a c re s quence, pr f rably a consensus seqeunce, and
additional amino acid residues that flank the core sequenc .
5 These peptides, including the methods for their
identification, are described in greater detail, below.
       Thus, in a specific embodiment of the invention,
peptides are provided which have at least nine and up to
about forty-five amino acid residues, including an amino acid
10 sequence resembling the formula,
       R-2-L-P-5-6-P-8-9 (SEQ ID NO:10),
positioned anywhere along the peptide. In the above-
mentioned formula, each number represents an amino acid
residue, such that 2 represents any amino acid residue except
15 cysteine, 5 and 6 each represents a hydrophobic amino acid
residue, 8 represents any amino acid residue except cysteine,
and 9 represents a hydrophilic amino acid residue except
cysteine. Each letter used in the formulas herein represent
the standard one-letter symbol for the corresponding amino
20 acid. When the peptide is a 9-mer, the peptide
R-P-L-P-P-L-P-T-S (SEQ ID NO:11) is excluded. The peptides
of particular interest are those that exhibit a binding
affinity for the SH3 domain of Src and Src-related proteins,
including Yes, Fyn, Lyn, Lck, Hck and Fgr. Preferably, the
25 peptides of the invention exhibit a binding affinity for the
SH3 domain of Src, which is at least three-fold, more
preferably at least four-fold, most preferably at least about
five-fold greater than that exhibited by the peptide RPLPPLP
(SEQ ID NO:9). In still other embodiments, the peptides
30 exhibit a binding affinity for the SH3 domain of Src which is
at least ten-fold greater than that exhibited by the peptide
RPLPPLP (SEQ ID NO:9).
       In specific embodiments, peptides are disclosed in which
the various amino acid residues at the indicated positions
35 may ind pendently have th following preferr d id ntities: 2
is a P, R, A, L, Q, E or S, more preferably P or R; 5
represents a P, M, I or L, more preferably P or M; 6 is a P,

- 14 -

SUBSTITUTE SHEET (RULE 26)

L, I or V, more preferably P or L; 8 is a T, R, P, I, N, E,
V, S, A, G r L, m r preferably T or R; and 9 is a T, R, S,
H or D, m r pr ferably T r R. Despit the preference for
hydrophobic amino acid residues at 5 and 6, in some cases it
5 may be desirable to have hydrophilic amino acid residues at
these positions. Specifically, amino acid residue 5 may be a
T, R or S, and amino acid residue 6 may be a T or R.
Likewise, while a hydrophilic amino acid residue is preferred
at position 9, in some instances a hydrophobic residue, such
10 as a P or A, may be desirable.

The present invention also contemplates SH3 domain-
binding peptides with a minimum length of 10, 11, 12, 13, 14,
15 or more amino acids. Such peptides contain additional
amino acid residues flanking the core sequence of
15 R-2-L-P-5-6-P (SEQ ID NO:71) either at the C-terminal end,
the N-terminal end or both. Thus, for example, such peptides
include those that further comprise a C-terminal-flanking
amino acid sequence of the formula 10, 10-11, 10-11-12, 10-
11-12-13 (SEQ ID NO:12) or 10-11-12-13-14 (SEQ ID NO:13), in
20 which each number represents any amino acid residue except
cysteine, such that the amino acid residue 10 is bound to the
amino acid residue 9 by a peptide bond. In that case,
specific embodiments include an amino acid residue 10 which
is T, R, L, S, D, P, A or N, preferably T or R, an amino acid
25 residue 11 which is R, P, A, Q, S or T, preferably R or P, an
amino acid residue 12 which is P, S, R or T, preferably P or
S, an amino acid residue 13 which is P, S, R, F, H or T,
preferably P or S, and an amino acid residue 14 which is S,
R, G or T, preferably, S or R.

30   Furthermore, peptides are also provided which further
comprise an N-terminal-flanking amino acid sequence of the
formula 1', 2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in
which each number represents any amino acid residue except
cysteine, such that 1' is bound to R by a peptide bond. In
35 such a cas , specific mbodiments are provided in which the
amino acid residu 1' is T, P, S, N, F, W, K, H, Q or G,
preferably T or P, wh rein the amino acid residue 2' is S, T,

- 15 -

G, P, R, Q, L, A or H, preferably S or T, wher in th amino
acid residu 3' is R, S, P, G, A, V, Y or L, preferably S r
T, and wherein the amin acid residu 4' is R, S, V, T, G, L
or F, preferably R or S.

5   In a particular embodiment, a peptide is disclosed
having at least thirteen and up to forty-five amino acid
residues, including an amino acid sequence of the formula,
3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15), positioned
anywhere along the peptide, in which each number represents
10 an amino acid residue, such that 3', 2', 1', 2, 8, and 10
each represents any amino acid residue except cysteine, 5 and
6 each represents a hydrophobic amino acid residue, and 9
represents a hydrophilic amino acid residue except cysteine,
each letter being the standard one-letter symbol for the
15 corresponding amino acid, said peptide exhibiting a binding
affinity for the SH3 domain of Src. Preferred 13-mers
include, but are not limited to, those having an amino acid
residue 5 which is a P or M, an amino acid residue 1' which
is T, P, S or N, an amino acid residue 2' which is S or T, an
20 amino acid residue 3' which is R or S, and an amino acid
residue 10 which is T or R.   In all the SH3 domain-binding
peptides described herein, the prohibition against the use of
the hydrophilic amino acid residue cysteine (C) does not
extend beyond the 7-mer "core" sequence and the additional
25 amino acid residues flanking the core up to a total (core +
flanking) of about 20 amino acids. That is, the occasional
use of a cysteine is not absolutely prohibited. What should
be kept in mind is that the potential for the formation of
intramolecular disulfide bonds, to form a cyclic structure,
30 be minimized as much as possible. Applicants have found that
cyclized structures appear to be disfavored, at least with
potential binding peptides of less than about 15 amino acid
residues in length. The concern for the formation of
cyclized structures comprising the core sequence diminishes
35 with increasing size of the peptide. Presumably, a large
en ugh structure, though cyclic, may allow the critical core
sequ nc to ad pt a m re or less lin ar conformation.

- 16 -

SUBSTITUTE SHEET (RULE 26)

In particular, specific p ptides ar disclos d which
exhibit binding affinities to SH3 d mains. These include th
peptides, RSTPRPLPMLPTTR (SEQ ID NO. 62), RSTPRPLPPLPTTR (SEQ
ID NO. 67), GILAPPVPPRNTR (SEQ ID NO. 63), VLKRPLPIPPVTR (SEQ
5 ID NO. 64), GPHRRLPPTPATR (SEQ ID NO. 65), and ANPSPATRPLPTR
(SEQ ID NO. 66).

Phage clones are also disclosed, along with the amino
acid sequences that are responsible for SH3 domain binding.
These phage clones are identified in Figure 5.

10   In other embodiments of the present invention, SH3
domain-binding peptides are contemplated which have a total
of 11, 13, 14, 18, 20, 22, 23, 25, 30, 36, 38 or 45 amino
acid residues.

The peptides of the present invention, having been
15 disclosed herein, may be prepared by any number of
practicable methods, including but not limited to solution-
phase synthesis, solid-phase synthesis, protein expression by
a transformed host, cleavage from a naturally-derived,
synthetic or semi-synthetic polypeptide, or a combination of
20 these techniques.

The SH3 binding peptides exhibit a wide range of
biological activity which includes the enhancement (or
inhibition, depending on the particular peptide or the natur
of the peptide's target molecule, in this case a protein
25 bearing an SH3 domain) of the natural function or biological
activity of the peptide's target molecule. For example, the
interaction of the binding peptide of the present invention
could result in the modulation of the oncogenic activity of
the target molecule bearing the SH3 domain. If the target
30 molecule has, in turn, a natural binding partner or ligand,
the peptides of the present invention may also exhibit
antagonistic or agonistic activity in relation to the
biological activity of the natural binding partner.

Thus, it is an object of the present invention to
35 provide a m thod of activating Src r Src-related pr t in
tyrosine kinas s by administering an eff ctive amount of the
SH3 domain-binding peptid s g nerally described herein. Th

- 17 -

SUBSTITUTE SHEET (RULE 26)

int nsity of the immune response can thus be stimulat d, f r
exampl , by th incr ased producti n of certain lymphokines,
such as TNF-alpha and interleukin-1. As is generally known
to those of ordinary skill in the art, a more intense immune
5 response may be in order in certain conditions, such as in
combating a particularly tenacious infection, viral or
otherwise, or a malignancy.

Furthermore, in a specific embodiment of the present
invention, a conjugate compound is contemplated which
10 comprises the peptide of the present invention and a second
chemical moiety. The second chemical moiety can be selected
from a wide variety of chemical compounds including the
peptide itself. Typically, however, the second chemical
moiety is selected to be other than the peptide of the
15 present invention, including but not limited to an amino
acid, a peptide other than an SH3 binding peptide of the
present invention, a polypeptide or protein (i.e., the
conjugate is a fusion protein), a nucleic acid, a nucleosid ,
a glycosidic residue (i.e., any sugar or carbohydrate), a
20 label or image-enhancing agent (including metals, isotopes,
radioisotopes, chromophores, fluorophores (such as FITC,
TRITC, and the like), and enzyme substrates), a drug
(including synthetic, semisynthetic, and naturally-occurring
compounds), small molecules (e.g., biotin, hormones, factors)
25 and the like.

The peptide of the present invention can be conjugated
to the second chemical moiety either directly (e.g., through
appropriate functional groups, such as an amine or carboxylic
acid group to form, for example, an amine, imine, amide,
30 ester, acyl or other carbon-carbon bond) or indirectly
through the intermediacy of a linker group (e.g., an
aliphatic or aromatic polyhydroxy, polyamine, polycarboxylic
acid, polyolefin or appropriate combinations thereof).
Moreover, the term "conjugate," as used herein, is also meant
35 to encompass non-covalent interacti ns, including but not
limited to ionic, affinity or other compl xation
interactions. Preferably, such other non-covalent

- 18 -

SUBSTITUTE SHEET (RULE 26)

interacti ns provid d finable, most pr ferably, is latable
chemical conjugate sp cies.

As described further herein, the peptides of th present
invention have been shown to localize within certain cellular
5 compartments which contain Src or Src-related proteins.
Consequently, the above-described conjugate can be utilized
as a delivery system for introduction of a drug to cells,
tissues or organs that include SH3 domain-containing
proteins.

10 It should also be pointed out that the present invention
seeks to provide a recombinant construct comprising a nucleic
acid or its complement that includes codons or nucleotide
sequences encoding a peptide having a region that binds to an
SH3 domain, preferably the Src SH3 domain. The recombinant
15 nucleic acid may be a DNA or RNA polynucleotide.

In a specific embodiment, the present invention
contemplates a recombinant construct which is a transforming
vector. Such vectors include those well known to those of
ordinary skill in the art, which effect the transfer or
20 expression of the nucleotide sequence after introduction to a
host, such as recombinant plasmid, phage or yeast artificial
chromosome. These vectors may be closed circular loops or
they may be linearized. The vectors contemplated include
those that exist extrachromosomally after host transformation
25 or transfection, as well as those that integrate within or
even displace portions of the host chromosome. The vectors
may be introduced to the cell with the help of transfection
aids or techniques well-known in the art. For example, these
aids or techniques may take the form of electroporation, us
30 of calcium chloride, calcium phosphate, DEAE dextran,
liposomes or polar lipid reagents known as LIPOFECTIN or
LIPOFECTAMINE. In addition, the present invention
contemplates the direct introduction of the desired nucleic
acid to the host cell, for instance, by injection.

35 Transformed h st c lls ar als obtained by the methods
of th present invention which ar capable of reproducing the
p lynucl otid s quences of int r st and/or expressing th

- 19 -

SUBSTITUTE SHEET (RULE 26)

c rresponding peptid products. A vari ty of h sts are
c ntemplat d, including prokary tic and eukaryotic hosts. In
particular, bacterial, viral, yeast, animal, and plant cells
are potentially transformable hosts. Thus, a method is

5 disclosed to obtain a transformed host cell that can produce,
preferably secrete, a peptide having a region that binds to
an SH3 domain comprising (a) providing an expression vector,
preferably a secretory expression vector, comprising a
nucleotide sequence encoding at least one copy of a peptide

10 having a region that binds to an SH3 domain; and (b)
introducing the vector to a competent host cell.

The peptides, thus produced, may then be introduced to
cells, tissues, organs, or administered to the subject for
the purpose of modulating the biochemical activity of the SH3

15 domain-containing proteins present therein. Accordingly, in
specific embodiments of the present invention, compositions
are provided which comprise an SH3 domain-binding peptide,
including a core sequence and flanking sequences, and a
suitable carrier.

20 The compositions contemplated by the present invention
may also include other components, from those that facilitate
the introduction or administration of the compositions to
those that have their own innate activity, such as a
prophylactic, a diagnostic or a therapeutic action. Such

25 innate activity may be distinct from that of the peptides of
the present invention or be complementary thereto. In any
event, the compositions of the present invention include
those that are suitable for administration into mammals,
including humans. Preferably, the compositions (including

30 necessarily the carrier) of the present invention are
sterile, though others may need only be cosmetically,
agriculturally or pharmaceutically acceptable. Still other
compositions may be adapted for veterinary use.

The compositions, including the drug delivery systems
35 d scribed h r in, are contemplat d to be administered in a
variety of ways, such as parent rally, orally, ent rally,
t pically or by inhalation. Th c mpositions may also be

- 20 -

SUBSTITUTE SHEET (RULE 26)

adminstered intranasally, pthalmically or intravaginally.
Furtherm r , the c mpositions f the inv nti n can tak
several forms, such as solids, gels, liquids, a r s ls or
patches.

5   In another embodiment of the present invention a method
is provided of identifying a peptide having a region that
binds to an SH3 domain comprising: (a) providing an
immobilized target protein comprising an SH3 domain; (b)
incubating the immobilized target protein with an aliquot
10 taken from a phage-displayed random peptide library, which
library includes peptides having a random sequence of ≥8
amino acid residues; (c) washing unbound phage from the
immobilized target protein; (d) recovering the phage bound to
the immobilized target protein; and (e) determining the
15 relevant nucleotide sequence of said binding phage nucleic
acid and deducing the primary sequence corresponding to the
SH3 domain-binding peptide. Preferably, the method further
comprises amplifying the titer of the recovered phage and
repeating the steps of incubation, washing and recovery to
20 provide SH3 domain-binding peptide-enriched phage.

Any other mode by which the peptide library, random or
otherwise, can be "displayed" can be utilized in the present
invention, however. Moreover, the present applicants believe
that longer *random* peptide sequences (e.g., >6 amino acid
25 residues, preferably >10, and most preferably, >12) provide
not only much greater diversity but also a richer degree of
secondary structure conducive to binding activity. If the
random region of the peptide is less than or equal to an 8-
mer, it should preferably not be cyclized.

30

5.2. Preparation of Random Peptide Libraries

The preparation and characterization of the
preferred phage-displayed random peptide libraries have been
described elsewhere. See, for example, Kay, B.K. et al. in
35 <u>Gene</u> (1992) 128:59-65, for a description of the preparation
of th phage-displayed rand m peptide library known as TSAR-
9, more below. In particular, by cloning deg n rat

- 21 -

SUBSTITUTE SHEET (RULE 26)

ligonucl otid s of fix d l ngth int bacteriophage vectors,
r combinant librari s f random peptides can be generated
which ar xpress d at th amino-t rminus of the pIII protein
on the surface of M13 viral particles. (There are 3-5 copies
5 of the pIII-fusion on the surface of each particle.) Phage
display offers several conveniences: first, the expressed
peptides are on the surface of the viral particles and
accessible for interactions; second, the recombinant viral
particles are stable (i.e., can be frozen, exposed to pH
10 extremes); third, the viruses can be amplified; and fourth,
each viral particle contains the DNA encoding the recombinant
genome. Consequently, these libraries can be screened by
isolating viral particles that bind to targets. The isolates
can be grown up overnight, and the displayed peptide sequence
15 responsible for binding can be deduced by DNA sequencing.

These libraries have approximately >10^8 different
recombinants, and nucleotide sequencing of the inserts
suggests that the expressed peptides are indeed random in
amino acid sequence. These libraries are referred to herein
20 as TSAR libraries, where TSAR stands for Totally Synthetic
Affinity Reagents. The preparation of the TSAR libraries are
described further below.

5.3. SH3 Binding Clones And Their Characteristics

25 Accordingly, peptides have been isolated from an
unconstrained random peptide library which exhibit a binding
affinity for SH3 domains. Furthermore, the binding
affinities exhibited by the disclosed peptides differ in
their selectivities with certain peptides showing comparable
30 binding affinities for SH3 domains derived from different
proteins, while others manifest greater affinities for
specific SH3 domains.

The amino acid sequence of various peptides isolated by
the present method are listed in Figure 5. As can be seen
35 from this list, certain groups of SH3 domain binding peptid s
are isolated fr m thre separat random p ptide libraries,
each based on a diff rent type of random peptide insert, all

- 22 -

SUBSTITUTE SHEET (RULE 26)

displayed at the amino-terminus of th pIII protein on th
surface of M13 viral particles. T n clon s w r isolated
from the R8C library, s v n fr m th TSAR-12 library, and
seven from the TSAR-9 library. The sequences are presented
5 to highlight the particular amino acid residues believed t
bind directly to the SH3 domain, as well as to point out th
remaining amino acid resiudes of the random insert and the
viral flanking sequences and complementary site amino acid
residues common to each group of clones. The frequency with
10 which each particular clone is found in each library is also
indicated in Figure 5. Thus, clones T12.SRC3.1 and
T12.SRC3.2 are by far the most abundant clones found among
the three libraries.

Interestingly, all the binding peptides are found to
15 have the proline-rich amino acid residue motif, which is
apparently responsible for binding, the motif being located
predominantly at the C-terminal end of the insert, although
each clone also contains an insert at the N-terminal end.
The significance of this observation is not presently
20 understood, although this finding may indicate the possible
importance of the C-terminal viral flanking sequences in SH3
domain binding.

Indeed, a synthetic peptide bearing only the core
consensus sequence RPLPPLP (SEQ ID NO:9) was less effective
25 in binding to target SH3 domains than synthetic peptides that
also included additional amino acid residues flanking the
core sequences. Thus, 13-mers and 14-mers having the
sequences RSTPRPLPMLPTTR (SEQ ID NO:62), RSTPRPLPPLPTTR (SEQ
ID NO:67), GILAPPVPPRNTR (SEQ ID NO:63), GPHRRLPPTPATR (SEQ
30 ID NO:65), and VLKRPLPIPPVTR (SEQ ID NO:64) have been
prepared and shown to bind to SH3 domains, such as those of
Src and Yes, much more avidly than the 7-mer, RPLPPLP (SEQ ID
NO:9). The 13-mer ANPSPATRPLPTR (SEQ ID NO:66) has been
shown to have binding affinities comparable to the core
35 consensus s qu nce. In each case, th 13-m rs compris a 7-
m r "core" s quence plus additional amino acid residu s

- 23 -

flanking sam , s me of which additional amino acid residues
are c ntributed by th viral flanking s quences.

Thus, in on embodim nt of the present inventi n, a 7-
mer core includes a consensus motif of the formula RXLPφφP
5 (SEQ ID NO:71), wherein R is arginine, L is leucine, P is
proline, X represents any amino acid except cysteine and φ
represents a hydrophobic amino acid residue. By "hydrophobic
amino acid residue," the applicants mean to include F, Y, W,
V, A, I, L, P or M, each letter representing the standard
10 one-letter designation for the corresponding amino acid
residue.

Furthermore, a preferred 9-mer peptide comprising two
additional amino acids on the C-terminal end of the core
sequence is envisioned having a consensus motif of the
15 formula RXLPφφPXψ. In this preferred 9-mer consensus motif,
the symbol ψ represents a hydrophilic amino acid residue,
except cysteine. By "hydrophilic amino acid residue," the
applicants mean to include K, R, H, D, E, N, Q, T, S or C,
and the other symbols are as defined above. For the purpos s
20 of the present invention, a glycine residue (G) may be
considered either a hydrophobic or a hydrophilic amino acid
residue. The one-letter symbols B and Z, which stand for N
or D and Q or E, respectively, are considered hydrophilic
amino acid residues.

25   Particular 13-mer peptides of the present invention
include those listed, below. It is noted, however, that not
all the following 13-mer peptides correlate strictly to or
comply with the preferred 9-mer consensus motif, described
above. Those peptides that *do not* comply (indicated in
30 italics, with the non-complying amino acid residues
underscored) can, thus, be described as "resembling" those
that do comply (indicated in normal type) with the preferred
9-mer consensus motif: PGFRELPPLPPSR (SEQ ID NO:72),
AQSRPLPIPPETR (SEQ ID NO:73), VLKRPLPIPPVTR (SEQ ID NO:64),
35 *PPNSPLPPLPTHL* (SEQ ID NO:74), *TGRGPLPPLPNDS* (SEQ ID NO:75),
YSTRPYPPITRPS (SEQ ID NO:76), *SHKSRLPPLPTRP* (SEQ ID NO:77),
YRFRALPSPPSAS (SEQ ID NO:78), *GPHRRLPPTPATR* (SEQ ID NO:65),

- 24 -

SUBSTITUTE SHEET (RULE 26)

LAQRQLPPTPGRD (SEQ ID NO:79), ALQRRLPRTPPPA (SEQ ID NO:80),
PATRPLPTRPSRT (SEQ ID NO:81), YSTRPLPSRPSRT (SEQ ID NO:82),
XPGRILLLPSEPR (SEQ ID NO:83), SGGILAPPVPPRN (SEQ ID NO:84),
RSTRPLPILPRTT (SEQ ID NO:85), STPRPLPMLPTTR (SEQ ID NO:86),
5 STNRPLPMIPTTR (SEQ ID NO:87), RSTRPLPSLPITT (SEQ ID NO:88),
STSRPLPSLPTTR (SEQ ID NO:89), RSTRSLPPLPPTT (SEQ ID NO:90),
RSTRQLPIPPTTT (SEQ ID NO:91), STPRPLPLIPTTP (SEQ ID NO:92),
RSTRPLPPTPLTT (SEQ ID NO:93), and RSTRPQPPPPITT (SEQ ID
NO:94). Accordingly, other peptides not specifically
10 disclosed, which either comply with or "resemble" the
preferred 9-mer consensus motif, can be readily envisioned by
those of ordinary skill in the art and are considered to b
equivalent to those that are specifically disclosed above.
In particular, non-compliance at positions 1 (S, G, and I, in
15 place of R, are tolerated), 3 (V, A, and Q, in place of L,
are tolerated), 4 (L, in place of P, is tolerated), 5
(hydrophilic amino acid residues, S, R, and T, are tolerat d
in place of a hydrophobic amino acid residue), 6 (hydrophilic
amino acid residues, R and T, are tolerated in place of a
20 hydrophobic amino acid residue), 7 (T, and S, in place of P,
are tolerated), and 9 (P and A are tolerated in place of a
hydrophilic amino acid residue) have been observed.

5.3.1.   Binding Specificities

25   It has been discovered that certain of the
binding peptides disclosed have a greater relative binding
affinity for one SH3 domain over another. Referring now to
Figure 8, the relative binding affinities of the various
peptides described above toward different SH3 domain targets
30 are graphically presented. As one can see, the relative
binding affinities of the respective peptides can differ by
orders of magnitude. Thus, as shown in Figure 8, the peptid
GPHRRLPPTPATR (SEQ ID NO:65), having the relevant sequence of
the phage clone identified as T12.SRC3.3, is specific to Src
35 family SH3 domains, including, but n t limited to, Src, Yes,
Lck, Hck, Fgr, Fyn, and Lyn. This SH3 binding peptide has
littl affinity for SH3 domains deriv d from PLCγ or Grb2.

- 25 -

SUBSTITUTE SHEET (RULE 26)

On the th r hand, th peptide GILAPPVPPRNTR (SEQ ID NO:63),
c rr sponding to the rel vant sequ nce of the phag cl n
T12.SRC3.1, which is one of the m st abundant binding clones
found by the present method, binds generically to a broad
5 range of SH3 domains, including Src, PLCγ, and Grb2.

On an intermediate level, the present invention has also
uncovered a peptide, VLKRPLPIPPVTR (SEQ ID NO:64),
corresponding to the relevant sequence of the phage clone
T12.SRC3.6, which is Src preferential; that is, this peptide
10 exhibits strong binding affinities for members of the Src
family, some binding affinities for Grb2 proteins, but littl
binding affinities for PLCγ domains. The peptide
ANPSPATRPLPTR (SEQ ID NO:66), corresponding to the relevant
sequence of the phage clone T12.SRC3.2, also exhibits Src
15 family specificity similar to GPHRRLPPTPATR (SEQ ID NO:65).
The peptides RSTPRPLPMLPTTR (R8C.YES3.5; SEQ ID NO:62) and
RSTPRPLPPLPTTR (representative consensus motif; SEQ ID NO:67)
are highly specific for SH3 domain of Src, Yes, and other
Src-related proteins.

20

5.4. Further Discussion of Binding Experiments

At the outset it is apparent that the binding
affinity of certain peptides to the SH3 domain of Src and
Src-related proteins is governed by more than just the
25 presence of the preferred core consensus sequences, RPLPPLP
(SEQ ID NO:9) or RPLPMLP (SEQ ID NO:95; i.e., RPLP(P/M)LP,
SEQ ID NO:96). Thus, while the synthetic peptides
RSTPRPLPMLPTTR (R8C.YES3.5; SEQ ID NO:62) and RSTPRPLPPLPTTR
(consensus; (SEQ ID NO:67) exhibit a strong specific binding
30 affinity for Src SH3, the other synthetic peptides tested
also exhibited an avid binding affinity to SH3 domains
relative to the 7-mer, RPLPPLP (SEQ ID NO:9). These other
peptides, GILAPPVPPRNTR (SEQ ID NO:63), VLKRPLPIPPVTR (SEQ ID
NO:64), GPHRRLPPTPATR (SEQ ID NO:65), and ANPSPATRPLPTR (SEQ
35 ID NO:66), sport cor sequ nc s and flanking s quenc s that
do n t losely adhere to the pref rred cor consensus
sequ nc s. Thus, th s results sugg st that binding affinity

- 26 -

SUBSTITUTE SHEET (RULE 26)

t SH3 domains is govern d to a larg ext nt by the nature of th amino acid residu s flanking th c re 7-m r sequ nce.

The binding charact ristics f Src SH3-s lected p ptides was determined using synthetic biotinylated peptides

5 corresponding to the sequences displayed by Src SH3-selected phage. These biotinylated peptides were assayed for direct binding to immobilized Src SH3-GST. Each of the five library-derived peptides tested were found to bind to Src SH3-GST and Yes SH3-GST over background (Figure 8).

10 Furthermore, a strong correlation was observed between the similarity of a given peptide to the preferred core consensus sequence RPLP(P/M)LP and the peptide's affinity for Src SH3-GST. The core sequence of the clone T12.SRC3.1 (GILAPPVPPRNTR; SEQ ID NO:63) appears to provide more generic

15 SH3 domain-binding characteristics.

Experiments comparing the relative binding of various phage clones to SH3 domains taken from a variety of proteins demonstrated the preference of these clones for Src and Src-related SH3 domains over SH3 domains taken from other

20 proteins.

It was further found that while the 7-mer having the consensus sequence RPLPPLP (SEQ ID NO:9) bound to Src SH3-GST only weakly, peptides comprising the consensus sequence flanked by residues encoded by one of the Src SH3-selected

25 clones (R8C.YES3.5), RSTP (SEQ ID NO:97) at the N-terminal end and TTR at the C-terminal end, bound significantly better than any of the peptides tested (Figure 7). Thus, as stat d previously, sequences that flank the RPLP(P/M)LP (SEQ ID NO:96) core appear to be important contributors to SH3

30 binding. It is further surmised that a peptide having or resembling the sequence RSTPAPPVPPRTTR (SEQ ID NO:98) should exhibit strong but generic binding to a variety of SH3 domains.

Similarly, it is observed that most of the Src SH3-

35 binding m tifs are located n ar the carboxy-t rminus of the random peptides, adjac nt t s quences which ar fix d in every clone (Figure 5). The exc ptional clones t nd to

- 27 -

SUBSTITUTE SHEET (RULE 26)

possess s quenc s that r semble motifs that include fixed
flanking s qu nces. This clustering contrasts with previous
results, in which binding m tifs ar distribut d throughout
the random peptide. Kay, B.K., et al., in _Gene_ (1993)
5 128:59-65.

The binding of the library-derived Src SH3-binding
peptides was compared to that of peptides corresponding to
proline-rich regions of natural proteins. Peptides
corresponding to SH3-binding regions in human PI-3' Kinase
10 (KISPPTPKPRPPRPLPV; SEQ ID NO:69) and human SOS1.20
(GTVEPVPPPVPPRRRPESA; SEQ ID NO:68), as well as a proline-
rich region of the cytoskeletal protein vinculin
(LAPPKPPLPEGEV; SEQ ID NO:70), bound Src SH3 less well than
the library-derived peptides (Figure 7).

15   As mentioned above, the relative specificity of binding
was explored. Thus, the relative binding of Src SH3-select d
peptides to equal amounts of GST fusions to SH3 domains from
different proteins was determined (Figure 8). While all of
the library-derived peptides bound the Src and Yes SH3
20 domains almost equally well, none of the peptides (with the
exception of peptide T12.SRC3.1, the most divergent peptide
tested) bound the SH3 domains of Grb2, Crk, Abl or PLCγ1
appreciably. Thus, the library-derived peptides, in contrast
with a peptide derived from SOS1, exhibit SH3 binding that is
25 relatively specific for Src-family members.

Next, it was determined whether the binding to the Src
SH3 domain was qualitatively like the interactions of the SH3
domain and natural proteins found in cell lysates. Thus,
radiolabeled proteins were prepared from NIH 3T3 cell lysates
30 and chromatographed over Src SH3-GST immobilized on
glutathione linked Sepharose. SDS-PAGE shows that a number
of proteins can be affinity purified in this manner. The
synthesized peptides bind quite well to the Src SH3 domain,
as they can compete the binding of radiolabeled proteins from
35 cell lysat s to imm biliz d Src-GST, with an IC_{50} of 1-10 mM
(Figure 9). In c nclusion, th p ptides can efficiently

- 28 -

SUBSTITUTE SHEET (RULE 26)

bl ck th interacti n f cellular proteins with Src SH3 *in vitro.*

M reover, *X nopus laevis* o cyt s inject d with mRNA encoding constitutively active Src undergo progesterone-

5 induced maturation at an accelerated rate relative to oocytes injected with water or c-Src mRNA. Unger, T.F. and Steele, R.E. in Mol. Cell.Biol. (1992) 12:5485-5498. To explore the ability of the library-derived Src SH3-binding peptides to exert a biochemical effect *in vivo*, the influence of the

10 peptides on the maturation of *Xenopus laevis* oocytes was examined. Hence, stage VI oocytes were injected with peptide, exposed to progesterone, and scored for germinal vesicle breakdown. Figure 10 shows that the rate of maturation was accelerated by approximately one hour when

15 oocytes were injected with the SH3-binding peptide consisting of RPLPPLP (SEQ ID NO:9) flanked by residues from clone T12.SRC3.6 (VLKRPLPIPPVTR; SEQ ID NO:64), but not with water or a peptide corresponding to a proline-rich segment of vinculin (LAPPKPPLPEGEV; SEQ ID NO:70) as controls. The

20 magnitude of this effect is roughly equivalent to that seen with injection of mRNA encoding constituitively active Src. See, e.g., Figure 3B in Unger, T.F. and Steele, R.E., *supra*. This result suggests that the library-derived Src SH3-binding peptide is effectively relieving an inhibitory effect of the

25 Src SH3 domain upon Src PTK activity. This model is consistent with a number of studies which have demonstrated an inhibitory effect of the Src SH3 domain upon Src kinase and transforming activity. See, e.g., Okada, M., et al., *supra*; Murphy, S.M., et al., *supra*; and Superti-Furga, G., et

30 al., *supra*.

5.5. Diagnostic And Therapeutic Agents
Based On SH3 Binding Peptides and Additional
Methods   of Their Use

As already indicated above, the present invention

35 also seeks to provide diagnostic, prophylactic, and

- 29 -

therapeutic agents bas d on the SH3 binding peptides
d scribed her in.

In one embodim nt, diagn stic agents ar provided,
preferably in the form of kits, comprising an SH3 domain-
5 binding peptide and a detectable label conjugated to said
peptide directly, indirectly or by complexation, said peptide
comprising: (i) a core sequence motif of the formula RXLPϕϕP
(SEQ ID NO:71), wherein X represents any amino acid except
cysteine and ϕ represents a hydrophobic amino acid residue,
10 including F, Y, W, V, A, I, L, P, M or G, each letter
representing the standard one-letter designation for the
corresponding amino acid residue; and (ii) two or more
additional amino acid residues flanking said core sequence at
its C-terminal end, N-terminal end or both.

15 The diagnostic agents of the present invention can be
used to detect the presence of SH3 domains of a generic or
specific type in cells, tissues or organs either *in vitro* or
in vivo. For *in vivo* applications, the diagnostic agent is
preferably mixed with a pharmaceutically acceptable carrier
20 for administration, either enterally, parenterally or by some
other route dictated by the needs of the particular
application.

In a particular embodiment, for example, an assay based
on a fusion product is contemplated which comprises a Src SH3
25 domain-binding peptide of the invention and a substrate for
deregulated or "activated" Src. For instance, a muscle
biopsy, taken from a subject suspected of being infected by
the Rous sarcoma virus, can be treated with an effective
amount of the fusion product. By subsequent analysis of the
30 degree of conversion of the substrate, one can potentially
detect infection by the Rous sarcoma virus in the subject,
particularly mammals, especially chickens. The presence of
the retrovirus, which causes the expression of deregulated or
"activated" Src, may thus be indicated by unusually high
35 levels of Src as rev aled by large amounts of th converted
substrate. S , for example, Paxton, W.G. et al., in
Biochem. Biophys. Res. Commun. (1994) 200(1):260-267

- 30 -

(detection f phosphorylated tyr sin and s rin residu s of
angiotensin II AT1 rec pt r, a substrate of Src family
tyrosine kinases); an ther suitabl substrat may b the
protein p68 (Fumagalli, S. et al., in Nature (1994)
5 368(6474):871-874; Taylor, S.J. and Shalloway, D., in Ibid.
at 867-871.

Alternatively, the enzyme can be isolated by selective
binding to a form of the SH3 domain-binding peptides of th
present invention (e.g., biotin-peptide conjugate). After
10 isolation of the protein-peptide conjugate complex (e.g., on
a column comprising streptavidin), the activity of the enzyme
can then be assayed by conventional methods to determine its
level of protein kinase activity which can be taken as an
indication of the presence of the deregulated or "activated"
15 form of the enzyme. An assay for Src kinase has been
described by Klinz and Maness, in Neuroprotocols (a companion
to Neuroscience)(1992) 1(3):224-231.

Moreover, the diagnostic agents of the invention can
also serve as imaging agents of cells, tissues or organs,
20 especially those that contain proteins with an SH3 domain.
For example, neural cells (e.g., neurons, other areas of th
brain), osteoclasts, osteoblasts, platelets, immune cells,
and other dividing cells are known to express or contain
proteins with SH3 domains. Thus, an image can be taken of
25 portions of the body to serve as a baseline for subsequent
images to detect physiologic or biochemical changes in the
subject's body. For instance, changes in the condition of
cellular levels of Src or a transformation of the cellular
Src to an "activated" form may be detected using the
30 diagnostic or imaging agents of the present invention.

Accordingly, it has been demonstrated that an SH3-
binding peptide tagged with a fluorescence emitter can
provide an image of the cytoskeleton. The images are
presented in Figure 11. As can be seen from Figure 11,
35 pan ls A, B, and C show th fluor sc nc image that is
obtained on tr ating NIH 3T3 fibroblasts with SH3 d main-
binding peptides modified t include a fluor scent tag. In

- 31 -

BNSDOCID <WO___9803649A1_IB>

sharp contrast, panel D shows only a dark image that is
produc d wh n th c lls ar tr ated with a prolin -rich
s gment f vinculin as a c ntr l.

In another embodiment, an SH3 domain-binding peptide-
5 horseradish immunoperoxidase complex or related
immunohistochemical agent could be used to detect and
quantitate specific receptor molecules in tissues, serum or
body fluids. In particular, the present invention provides
useful diagnostic reagents for use in immunoassays, Southern
10 or Northern hybridization, and *in situ* assays. Accordingly,
the diagnostic agents described herein may be suitable for
use *in vitro* or *in vivo*.

In addition, the diagnostic or imaging agent of the
present invention is not limited by the nature of the
15 detectable label. Hence, the diagnostic agent may contain
one or more such labels including, but not limited to,
radioisotope, fluorescent tags, paramagnetic substances,
heavy metals, or other image-enhancing agents. Those of
ordinary skill in the art would be familiar with the range f
20 label and methods to incorporate or conjugate them into the
SH3 domain-binding peptide to form diagnostic agents.

In yet a further embodiment, pharmaceutical compositions
are provided comprising an SH3 domain-binding peptide and a
pharmaceutically acceptable carrier. In a specific
25 embodiment of the invention, the pharmaceutical composition
is useful for the modulation of the activity of SH3 domain-
containing proteins. By "modulation" is meant either
inhibition or enhancement of the activity of the protein
target. Accordingly, a pharmaceutical composition is
30 disclosed comprising an SH3 domain-binding peptide and a
pharmaceutically acceptable carrier, said peptide comprising:
(i) a 9-mer sequence motif of the formula RXLP$\phi\phi$PXψ (SEQ ID
NO:10), wherein X represents any amino acid except cysteine,
ϕ represents a hydrophobic amino acid residue, and wherein ψ
35 is a hydr philic amino acid r sidu xcept cysteine, each
l tter r pr s nting th standard on -letter designation for
the corr sponding amino acid residue; and, optionally, (ii)

- 32 -

additi nal amino acid residues flanking th 9-m r s quence at
its C-t rminal end, N-terminal end or both, up to a total f
45 amino acid residues, including said 9-m r s quenc .
Preferably, the peptide comprises at least one, more

5 preferably at least two, and most preferably at least thre
additional amino acids flanking the 9-mer sequence.

As stated above, the therapeutic or diagnostic agents f
the invention may also contain appropriate pharmaceutically
acceptable carriers, diluents and adjuvants. Such

10 pharmaceutical carriers can be sterile liquids, such as water
and oils including those of petroleum, animal, vegetable or
synthetic origin, such as peanut oil, soybean oil, mineral
oil, sesame oil and the like. Water is a preferred carrier
when the pharmaceutical composition is administered

15 intravenously. Saline solutions and aqueous dextrose and
glycerol solutions can also be employed as liquid carriers,
particularly for injectable solutions. Suitable
pharmaceutical excipients include starch, glucose, lactose,
sucrose, gelatin, malt, rice, flour, chalk, silica gel,

20 magnesium carbonate, magnesium stearate, sodium stearate,
glycerol monostearate, talc, sodium chloride, dried skim
milk, glycerol, propylene, glycol, water, ethanol and the
like. These compositions can take the form of solutions,
suspensions, tablets, pills, capsules, powders, sustained-

25 release formulations and the like. Suitable pharmaceutical
carriers are described in "Remington's Pharmaceutical
Sciences" by E.W. Martin.

Such compositions will contain an effective therapeutic
amount of the active compound together with a suitable amount

30 of carrier so as to provide the form for proper
administration to the subject. While intravenous injection
is a very effective form of administration, other modes can
be employed, including but not limited to intramuscular,
intraperitoneal, and subcutaneous injection, and oral, nasal,

35 enteral, and par nteral administration.

The th rap utic agents and diagnostic agents of the
instant inv ntion are us d f r th treatment and/or diagnosis

- 33 -

SUBSTITUTE SHEET (RULE 26)

of animals, and mor preferably, mammals including humans, as
well as dogs, cats, horses, cows, pigs, guin a pigs, mic and
rats. Accordingly, ther methods c ntemplated in th pres nt
invention, include, but are not limited to, a method of

5 modulating, i.e., inhibiting or enhancing, bone resorption in
a mammal (see, e.g., Hall, T.J., in Biochem. Biophys. Res.
Commun. (1994) 199(3):1237-44), a method of disrupting
protein tyrosine kinase-mediated signal transduction pathways
or a method of regulating the processing, trafficking or

10 translation of RNA in a cell by introducing or administering
an effective amount of an SH3 domain-binding peptide of the
present invention (see, e.g., Taylor, S.J. and Shalloway, D.,
supra).

     The diagnostic or therapeutic agents of the present

15 invention can be modified by attachment to soluble
macromolecules such as proteins, polysaccharides, or
synthetic polymers. For example, the peptide could be
coupled to styrene-maleic acid copolymers (see, e.g.,
Matsumura and Maeda, Cancer Res. (1986) 46:6387),

20 methacrylamide copolymers (Kopececk and Duncan, J. Controlled
Release (1987) 6:315), or polyethylene glycol (PEG) (e.g.,
Hershfield and Buckley, N. Engl. J. Med. (1987) 316:589; Ho
et al., Drug Metab. Dispos. (1986) 14:349; Chua et al., Ann.
Intern. Med. (1988) 109:114).   The agents, if desired,

25 are further targeted by attachment to an antibody, especially
a monoclonal antibody. Such antibodies include but are not
limited to chimeric, single chain, Fab fragments, and Fab
expression libraries. In one embodiment the agent is coupl d
to the macromolecule via a degradable linkage so that it will

30 be released in vivo in its active form.

     In another embodiment, the therapeutic or diagnostic
agent may be delivered in a vesicle, in particular a
liposome. See, Langer, Science (1990) 249:1527-1533; Treat
et al., in Liposomes in the Therapy of Infectious Disease

35 and Cancer, Lop z-Berestein and Fidler (eds.), Liss, New York
(1989) pp. 353-365; Lopez-Berestein, ibid., pp. 317-327.

- 34 -

In yet another embodiment, the therapeutic or *in vivo* diagnostic agent can be delivered in a controlled release system. In one embodiment, a pump may be used (se' Langer, *supra*; Sefton, <u>CRC Crit. Ref. Biomed. Eng.</u> (1987) 14:201;

5 Buchwald et al., <u>Surgery</u> (1980) 88:507; Saudek et al., <u>N.</u> <u>Engl. J. Med.</u> (1989) 321:574). In another embodiment, polymeric materials may be used (see Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida, 1974; Controlled Drug Bioavailability, Drug

10 Product Design and Performance, Smolen and Ball (eds.) Wiley, New York 1984; Raner and Peppas, <u>J. Macromol. Sci. Rev.</u> <u>Macromol. Chem.</u> (1983) 23:61; see, also, Levy et al., <u>Science</u> (1985) 228:190; During et al., <u>Ann. Neurol.</u> (1989) 25:351; Howard et al., <u>J. Neurosurg.</u> (1989) 71:105). In a preferred

15 embodiment, a controlled release system may be placed next to the therapeutic target, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in <u>Medical Applications of</u> <u>Controlled Release</u>, *supra*, (1984) 2:115-138). It will be recognized by one of ordinary skill in the art that a

20 particular advantage of the invention is that a peptide will not be subject to the problems of denaturation and aggregation associated with proteins held in the warm, most environment of a body in a controlled release system.

Other controlled release systems are discussed in the

25 review by Langer, in <u>Science</u> (1990) 249:1527-1533.

30

35

SUBSTITUTE SHEET (RULE 26)

6.   **EXAMPLES**

6.1. **Preparati n f the TSAR-9 Library**

6.1.1.   **Synthesis and Assembly f Oligonucleotides**

Figure 1 shows the formula of the oligonucleotides and the assembly scheme used in construction of the TSAR-9 library. The oligonucleotides were synthesized with an applied Biosystems 380a synthesizer (Foster City, CA), and the full-length oligonucleotides were purified by HPLC.

Five micrograms of each of the pair of oligonucleotides were mixed together in buffer (10 mM Tris-HCl, pH 8.3, 15 mM KCl, 0.001% gelatin, 1.5 mM magnesium chloride), with 0.1 % Triton X-100, 2 mM dNTP's, and 20 units of *Taq* DNA polymerase. The assembly reaction mixtures were incubated at 72 °C for 30 seconds and then 30 °C for 30 seconds; this cycle was repeated 60 times. It should be noted that the assembly reaction is not PCR, since a denaturation step was not used. Fill-in reactions were carried out in a thermal cycling, device (Ericomp, LaJolla, CA) with the following protocol: 30 seconds at 72 °C, 30 seconds at 30 °C, repeated for 60 cycles. The lower temperature allows for annealing of the six base complementary region between the two sets of the oligonucleotide pairs. The reaction products were phenol/chloroform extracted and ethanol precipitated. Greater than 90% of the nucleotides were found to have been converted to double stranded synthetic oligonucleotides.

After resuspension in 300 μL of buffer containing 10 mM Tris-HCI, pH 7.5, 1 mM EDTA (TE buffer), the ends of the oligonucleotide fragments were cleaved with *Xba* I and *Xho* I (New England BioLabs, Beverly, MA) according to the supplier's recommendations. The fragments were purified by 4% agarose gel electrophoresis. The band of correct size was removed and electroeluted, concentrated by ethanol precipitation and resuspended in 100 μL TE buffer. Approximately 5% of the assembled oligonucl otides can be expected to have int rnal *Xho* I or *Xba* I sites; how v r, only

- 36 -

BNSDOCID: <WO__9603649A1_IB>

th full-length m lecul s wer us d in the ligation step of th assembly scheme. The conc ntration of the synth tic ligonucl tide fragm nts was estimated by comparing th intensity on an ethidium bromide stained gel run along with

5 appropriate quantitated markers. All DNA manipulations not described in detail were performed according to Maniatis, *supra*.

To demonstrate that the assembled enzyme digested oligonucleotides could be ligated, the synthesized DNA

10 fragments were examined for their ability to self-ligate. The digested fragments were incubated overnight at 18 °C in ligation buffer with T4 DNA ligase. When the ligation products were examined by agarose gel electrophoresis, a concatamer of bands was visible upon ethidium bromide

15 staining. As many as five different unit length concatamer bands (i.e., dimer, trimer, tetramer, pentamer, hexamer) w re evident, suggesting that the synthesized DNA fragments wer efficient substrates for ligation.

20   6.1.2.   Construction of Vectors

The construction of the M13 derived phage vectors useful for expressing a TSAR library has been recently described (Fowlkes, D. et al. BioTech. (1992) 13:422-427). To express the TSAR-9 library, an M13 derived

25 vector, m663, was constructed as described in Fowlkes. Th m663 vector contains the pIII gene having a *c-myc*-epitope, i.e., as a stuffer fragment, introduced at the mature N-terminal end, flanked by *Xho* I and *Xba* I restriction sites (see also, Figure I of Fowlkes).

30

      6.1.3.   Expression of the TSAR-9 Library

The synthesized oligonucleotides were then ligated to *Xho* I and *Xba* I double-digested m663 RF DNA containing, the pIII gene (Fowlkes) by incubation with ligas

35 v rnight at 12 °C. Mor particularly, 50 ng of v ctor DNA and 5 ng of th digest d synth sized DNA and was mixed tog ther in 50 μL ligation buff r (50 mM Tris, pH 8.0, 10 mM

- 37 -

MgCl$_2$, 20 mM DTT, 0.1 mM ATP) with T4 DNA ligas . After
ov rnight ligati n at 12°C, the DNA was concentrated by
ethanol precipitation and washed with 70% ethanol. The
ligated DNA was then introduced into *E. coli* (DH5αF'; GIBCO
5 BRL, Gaithersburg, MD) by electroporation.

A small aliquot of the electroporated cells was plated
and the number of plaques counted to determine that 10⁸
recombinants were generated. The library of *E. coli* cells
containing recombinant vectors was plated at a high density
10 (~400,000 per 150 mM petri plate) for a single amplification
of the recombinant phage. After 8 hr, the recombinant
bacteriophage were recovered by washing each plate for 18 hr
with SMG buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM
MgCl$_2$, 0.05% gelatin) and after the addition of glycerol to
15 50% were frozen at -80 °C. The TSAR-9 library thus formed
had a working titer of ~2 x 10¹¹ pfu/ml.

6.2. Preparation of TSAR-12 Library

Figure 2 shows the formula for the synthetic
20 oligonucleotides and the assembly scheme used in the
construction of the TSAR-12 library. As shown in Figure 2,
the TSAR-12 library was prepared substantially the same as
the TSAR-9 library described in Section 6.1 above with the
following exceptions: (1) each of the variant non-predicted
25 oligonucleotide sequences, i.e., NNB, was 30 nucleotides in
length, rather than 54 nucleotides; (2) the restriction sites
included at the 5' termini of the variant, non-predicted
sequences were *Sal* I and *Spe* 1, rather than *Xho* I and *Xba* I;
and (3) the invariant sequence at the 3' termini to aid
30 annealing of the two strands was GCGGTG and CGCCAC rather
than CCAGGT and GGTCCA (5' to 3').

After synthesis including numerous rounds of annealing
and chain extension in the presence of dNTP's and Taq DNA
polymerase, and purification as described above in Section
35 6.1.1, the synth tic d uble stranded, oligonucleotide
fragments were dig sted with *Sal* I and *Spe* I restriction
enzymes and ligat d with T4 DNA ligas t the nucl tide

- 38 -

SUBSTITUTE SHEET (RULE 26)

sequence encoding the M13 pIII gene contained in the m663
vector to yield a library f TSAR- xpr ssion vect rs as
described in S cti ns 6.1.2 and 6.1.3. Th ligat d DNA was
then introduced into *E. coli* (DH5αF'; GIBCO BRL,

5 Gaithersburg, MD by electroporation. The library of *E. coli*
cells were plated at high density (~400,000 per 150 mm petri
plate) for amplification of the recombinant phage. After
about 8 hr, the recombinant bacteriophage were recovered by
washing, for 18 hr with SMG buffer and after the addition of

10 glycerol to 50% were frozen at -80 °C.

The TSAR-12 library thus formed had a working titer of
~2 x 10^{11} pfu/mL.

6.3. Characterization of the TSAR-9 and -12 Libraries

15   The inserted synthetic oligonucleotides for each of
the TSAR libraries, described in Sections 6.1 and 6.2 above,
had a potential coding complexity of 20^{36} (~10^{47}) and 20^{20},
respectively, and since ~10^{14} molecules were used in each
transformation experiment, each member of these TSAR

20 libraries should be unique. After plate amplification the
library solution or stock has 10^4 copies of each member/mL.

It was observed that very few (<10%) of the inserted
oligonucleotide sequences characterized so far in both of th
libraries have exhibited deletions or insertions. This is

25 likely a reflection of the accuracy assembling the
oligonucleotides under the conditions used and the fact that
certain types of mutations (i.e., frame-shifts) would not be
tolerated as pIII an essential protein for phage propagation.

In order to determine whether any coding bias existed in

30 the variant non-predicted peptides expressed by these
libraries, perhaps due to biases imposed *in vitro* during
synthesis of the oligonucleotides or *in vivo* during
expression by the reproducing phage, inserts were sequenced
as set forth below.

35

6.3.1.   Charact rizati n f TSAR-9 Library

- 39 -

Insert d synthetic ligonucleotid fragments
of 23 rand mly chosen is lates wer examined fr m th TSAR-9
library. Individual plaqu s w r us d to in culate I ml f
2XYT broth containing *E. coli* (DH5αF') cells and the cultures
5 were allowed to grow overnight at 37°C with aeration. DNA
was isolated from the culture supernatants according to
Maniatis, *supra*. Twenty-three individual isolates were
sequenced according to the method of Sanger (*Proc. Natl.
Acad. Sci. USA* (1979) 74:5463-5467) using as a primer the
10 oligonucleotide 5'-AGCGTAACGATCTCCCG (SEQ ID NO. 99), which
is 89 nucleotides downstream of the pIII gene cloning site of
the m663 vector used to express the TSARS.

Nucleotide sequences and their encoded amino acid
sequences were analyzed with the MacVector computer program
15 (IBI, New Haven, CT). The Microsoft EXCEL program was used
to evaluate amino acid frequencies. Such analyses showed
that the nucleotide codons coding for and hence most amino
acids, occurred at the expected frequency in the TSAR-9
library of expressed proteins. The notable exceptions were
20 glutamine and tryptophan, which were over- and under-
represented, respectively.

It is of interest to note the paucity of TAG stop codons
in the inserts, i.e., only 2 of -200 isolates characterized
contained a TAG stop codon. About half $[1-(47/48)^{36}]$ of the
25 phage inserts were expected to have at least one TAG codon in
view of the assembly scheme used. However, most of the TAG-
bearing phage appear to have been lost from the library, even
though the bacterial host was *supE*. This may be a
consequence of suppression being less than 100% effective.
30 The amino acids encoded by the inserted double stranded
synthesized oligonucleotide sequences, excluding the fixed
PG-encoding centers, were concatenated into a single sequence
and the usage frequency determined for each amino acid using
the Microsoft EXCEL program. These frequencies were compared
35 to that expected from the ass mbly schem of the
oligonucleotides, and th div rg nce fr m exp cted values
represented by the size of the bars above and below th

- 40 -

SUBSTITUTE SHEET (RULE 26)

bas line. Chi squar analysis was used to det rmin the
significance of the d viations. Th majority of amino acids
w r f und to occur at th expect d fr quency, with th
notable exceptions that glutamine and tryptophan were
5 somewhat over- and under-represented, respectively. Thus,
except for the invariant Pro-Gly, any position could have any
amino acid; hence, the sequences are unpredicted or random.

6.3.2.   Characterisation of TSAR-12 Library

10   Approximately 10 randomly chosen inserted
oligonucleotides from the TSAR-12 library were examined by
DNA sequencing as described above in Section 6.3.1. The
isolates were chosen at random from the TSAR-12 library and
prepared for sequencing, as were the TSAR-9 isolates.
15 Analysis showed that except for the invariant Gly any
position could have any amino acid; hence, the sequences ar
unpredicted or random.

6.4. Preparation of R8C Library

20   Referring now to Figure 3, two oligonucleotides
were synthesized on an Applied Biosystems Model 380a machine
with the sequence 5'-
TGACGTCTCGAGTTGTNNKNNKNNKNNKNNKNNKNNKNNKTGTGGATCTAGAAGGATC-3'
(SEQ ID NO:31) and 5'-GATCCTTCTAGATCC-3' (SEQ ID NO:32),
25 where N is an equimolar ratio of deoxynucleotides A, C, G,
and T, and K is an equimolar ratio of G and T. Fifty pmol of
each oligonucleotide was incubated at 42 °C for 5 min, then
37 °C for 15 min, in 50 μL of Sequenase™ buffer (U.S.
Biochemicals, Cleveland, OH) with 0.1 μg/μL acetylated BSA,
30 and 10 mM DTT. After annealing, 10 units of Sequenase™ (U.S.
Biochemicals) and 0.2 mM of each dNTP were added and
incubated at 37 °C for 15 min. The sample was then heated at
65 °C for 2 hr, digested with 100 units of both *Xho* I and *Xba*
I (New England BioLabs, Beverly, MA), phenol extracted,
35 thanol pr cipitat d, and resolved on a 15% non-denaturing
polyacrylamide g l. Th ass mbl d, digested fragment was gel
purified pri r to ligation. Th vector, m663 (Fowlkes, D. t

- 41 -

al. <u>Biotech</u>. (1992) 13:422-427), was pr pared by digestion
with *Xho* I and *Xba* I, calf alkalin ph sphatas (Boehringer
Mannheim, Indianap lis, IN) tr atment, ph nol extracted, and
purified by agarose gel electrophoresis. To ligate, 20 μg
5 vector was combined with 0.2 μg insert in 3 mL with T4 DNA
ligase (Boehringer Mannheim), according to the manufacturer.
After removal of the protein and buffer by phenol extraction
and ethanol precipitation, the ligated DNA was electroporat d
into XL1-Blue *E. coli* (Stratagene, San Diego, CA) and plated
10 for eight hours at 37 °C. To recover the recombinant phage,
the top agar was collected with a spatula, mixed with an
equal volume of 100 mM NaCl, 10 mM $MgCl_2$, and 50 mM Tris-HCI
(pH7.5), and disrupted by two passes through an 18-gauge
syringe needle. The bacterial cells were removed by
15 centrifugation, and phage particles were collected by
polyethylene glycol precipitation and stored at -70 °C in 25%
glycerol. The library had 10^8 total recombinants and a
working titer of 6 x 10^{13}pfu/mL.

Members of the library were checked for inserts by the
20 polymerase chain reaction (Saiki, et al. <u>Science</u> (1988)
239:487-491). Individual plaques on a petri plate were
touched with a sterile toothpick and the tip was stirred int
2xYT with F$^+$ *E. coli* bacteria and incubated overnight at 37 °C
with aeration. Five microliters of the phage supernatant
25 were then transferred to new tubes containing buffer (67 mM
Tris-HCl, pH 8.8/10 mM β- mercaptoethanol/16.6 mM ammonium
sulfate/6.7 mM EDTA/50 μg bovine serum albumin per mL), 0.1
mM deoxynucleotide triphosphates, and 1.25 units of *Taq* DNA
polymerase (Boehringer Mannheim, Indianapolis, IN) with 100
30 pmoles of oligonucleotide primers. The primers flanked the
cloning site in gene III of m663 (5'-TTCACCTCGAAAGCAAGCTG-3'
(SEQ ID NO:100) and
5'-CCTCATAGTTAGCGTAACG-3' (SEQ ID NO:101)). The assembly
reactions were incubated at 94 °C for 1 min, 56 °C for 2 min,
35 and 72 °C for 3 min; this cycl was r p at d 24 times. The
reaction products w r th n resolved by electrophoresis on a
NuSi ve 2.0% agaros g l (FMC, Rockland, ME). Gels rev aled

- 42 -

SUBSTITUTE SHEET (RULE 26)

that for 20 plaqu s test d, all w re r combinant and had
singl inserts of th   xpected siz .

Based on the sample size f the library, it was
anticipated that 100% of the recombinants had single inserts.
5 However, all of the SH3-binding phage isolated from the R8C
library had double-inserts. Such phage are presumed rare
(i.e., <5%) within the library, yet because the SH3-binding
peptide appears to need to be linear they were selected for
by our screening methods. Most likely they were formed
10 during the generation of the library; one scenario is that
the inserts ligated together to form head-to-head dimers and
that they were subsequently cloned into m663 DNA by ligation
with the vector's *Xho I* sticky end and by illegitimate
ligation with the vector's *Xba I* site (see, Figure 4).

15

6.5. Preparation Of Target-Coated Microtiter Wells

6.5.1.   Preparation Of GST-SH3 Fusion Proteins

The preparation of Src-GST fusion protein
was first described by Smith and Johnson, in <u>Gene</u> (1988)
20 67:31, the disclosure of which is incorporated by reference
herein. Briefly, pGEX-derived (Pharmacia, Piscataway, NJ)
constructs expressing GST fusion proteins containing the SH3
domains of Src, Grb2, Crk, Abl, or PLCγ were obtained from
Dr. Channing Der (University of North Carolina at Chapel
25 Hill); a construct expressing the SH3 domain of Yes was
obtained from Dr. Marius Sudol (Rockefeller University). The
use of the pGEX bacterial expression vector for the
production of GST-SH3 fusion proteins is well-known to thos
in the art. See, e.g., Cicchetti, P. et al., in <u>Science</u>
30 (1992) 257:803-806. Briefly, the coding region for a
particular SH3 domain can be fused in-frame at the Bam HI
site of pGEX-2T. Thus, fusion proteins were prepared as per
the manufacturer's instructions, and quantified by Coomassie
Blue staining of SDS-polyacrylamide gels. Microtiter wells
35 wer coated with 5-20 μg GST-SH3 fusion protein in 100 mM
$NaHCO_3$, pH 8.5, blocked with 100 mM $NaHCO_3$ (pH 8.5) 1% BSA,
and washed. All washes consisted of five applications of

- 43 -

1XPBS, 0.1% Tween 20, 0.1% BSA (Buff r A). Where
appropriat , th am unt of protein bound t ach w ll was
quantifi d with an anti-GST antibody-based ELISA (Pharmacia,
Piscataway, NJ), and with a GST-binding phage, isolated
5 during the course of this work.

6.5.2.   Coating of Microtiter Wells

        Bacterially expressed Src SH3 glutathione-S-
transferase (Src-GST) fusion protein was purified from
bacterial lysates using glutathione agarose 4B (Pharmacia),
10 according to the manufacturer's instructions. Bound Src-GST
fusion protein was eluted from the glutathione agarose with
10 mM glutathione in PBS. Microtiter wells were then coated
with Src-GST fusion protein (1-10 μg/well, in 50 mM NaHCO$_3$, pH
8.5) overnight at 4 °C. To block non-specific binding of
15 phage, 100 μL 1% BSA in 100 mM NaHCO$_3$, pH 8.5, was added to
each well and allowed to incubate at room temperature for 1
hour. The wells were then washed five times with 200 μL PBS,
0.1% Tween 20, 0.1% BSA (Buffer A).

20   6.6. Biopanning And Subsequent Characterization Of
       Phage-Displayed Random Peptide Libraries
       With Src-GST Fusion Protein As Target Molecule

6.6.1.   Isolation of Src SH3-Binding Phage

        Library screens were performed as
previously described. Kay, B.K., et al., in Gene (1993)
25 128:59-65. Briefly, 1 X 10^{11} pfu TSAR 9, TSAR 12, or R8C
phage in Buffer A were incubated in a Src SH3-GST-coated well
for 2 hours. The wells were washed, and bound phage were
eluted with 100 μL 50 mM glycine·HCl (pH 2.2), transferred to
a new well, and neutralized with 100 mL 200 mM NaHPO$_4$ (pH
30 7.0). Recovered phage were used to infect 1 x 10^9 DH5αF' E.
coli cells in 20 mL 2xYT; the infected cells were grown
overnight, resulting in a 1000- to 10,000-fold amplification
of phage titer. Amplified phage were panned twice more, as
above, excepting the amplification step. Binding phage
35 recover d after the third round of panning were plat d at a
low density on a lawn f DH5αF' E. coli c lls to yield

- 44 -

SUBSTITUTE SHEET (RULE 26)

is lated plaqu s f r cl nal analysis. Isolated plaques were
us d t produc small cultures fr m which phage stocks and
DNA were r cover d for phag binding experiments and dide xy
sequencing (Sanger, F., et al., in <u>Proc. Natl. Acad. Sci. USA</u>
5 (1977) 74:5463-5467), respectively. Clones were confirmed as
binding the SH3 domain by applying equal titers of phage to
wells containing Src SH3-GST or GST alone, and titering the
number of eluted particles from each well, or detecting bound
phage with an anti-phage antibody-based ELISA (Pharmacia).

10   Indeed, the ability of isolated phage clones to bind to
several SH3 domains derived from a variety of different
proteins can be investigated by the manner described above.
GST-SH3 fusion proteins containing SH3 domains from a variety
of different proteins are bound to microliter wells. An
15 aliquot of the aforementioned phage stocks (50 μL) is
introduced into wells containing the different GST-SH3 fusion
proteins. After room temperature incubation for 1-2 hours,
the liquid contents of the microtiter plates are removed, and
the wells are washed 5 times with 200 μL Buffer A. Bound
20 phage are eluted with 100 μL 50 mM glycine (pH 2.2),
transferred to a new well, and neutralized with 100 μL 200 mM
NaHPO$_4$ (pH 7.0). The phage are diluted 10^{-3}- to 10^{-6}-fold, and
aliquots are plated onto lawns of DH5αF' $E. coli$ cells to
establish the number of plaque forming units in the output
25 sample. From these experiments, the relative specificity of
different Src SH3 binding clones for SH3 domains derived from
other proteins is determined.

        6.6.2.   Phage ELISA and Nucleotide Sequencing
30   To evaluate the binding of isolates to
various targets proteins, enzyme-linked-immuno-assays (ELISA)
were also performed. Bacterial cultures were infected with
phage isolates and cultured overnight in 2XYT at 37 °C. The
cells were spun down and 25 mL of supernatant was added to
35 microtiter plate w lls c ated with 50 mL of protein (1 mg/mL
in 100 mM NaHCO$_3$, pH 8.4; overnight at 4 °C or f r a few hours
at room temperature) and blocked (1 mg/mL BSA in 100 mM

- 45 -

SUBSTITUTE SHEET (RULE 26)

NaHCO$_3$, pH 8.4; f r about one hour). The phag are incubat d
in the w ll with 25 mL of PBS-0.1% Tween 20 at RT for 2 hr.
The wells are then washed multipl times over 30 minutes. To
each well is added 50 μL of polyclonal anti-phage antibody
5 conjugated to horseradish peroxidase. The antibody is
diluted 1:3000 in BSA-PBS-Tween 20; it was obtained from
Pharmacia (Piscataway, NJ; catalog number 27-9402-01). Aft r
30 minutes, the wells are washed again with BSA-PBS-Tween 20
for ¯20 minutes. Finally, 100 μL of ABTS reagent (Pharmacia,
10 with H$_2$O$_2$) are added to each well for the development of
color. Plates are read with a plate reader (Molecular
Devices, Menlo Park, CA) at 405 nm wavelength.

The nucleotide sequence of the relevant segments of th
Src SH3 binding clones (or phage clones that bind to SH3
15 domains of other proteins) were sequenced using standard
methods. Sanger, F., et al., in _Proc. Natl. Acad. Sci. USA_
(1977) 74:5463-5467. The oligo primer 5'-AGCGTAACGATCTAAA-3'
(SEQ ID NO:102) was used, which is 89 nucleotides downstream
of the gene _III_ cloning site of M13 m666. The nucleotide
20 sequences were analyzed with the MacVector computer program
(IBI, New Haven, CT, USA). From this nucleotide sequence
information the primary sequence of each Src SH3 binding
peptide was deduced. The corresponding synthetic peptides
were then prepared by techniques well known in the art with
25 or without flanking sequences. Indeed, these synthetic
peptides have been shown to bind to SH3 domain targets, with
those possessing the phage flanking amino acid residues
exhibiting greater binding affinity.

30   **6.7   _In Vitro_ Peptide Binding Assays**
Peptides were obtained from Research Genetics
(Birmingham, AL), Chiron Mimotopes (Victoria, Australia), or
synthesized by conventional techniques by Dr. J. Mark Carter
of Cytogen Corporation (Princeton, NJ). Peptide purity was
35 assess d by HPLC and/ r mass sp ctrom try. Biotinylat d
peptid s were synthesized with either a KSGSG (SEQ ID NO:103)
or a GSGS (SEQ ID NO:104) peptid link r (a spac r) betw en

- 46 -

SUBSTITUTE SHEET (RULE 26)

the bi tin and th N-terminus of the peptide. Binding
experiments wer p rformed as ab v , excepting th use of 10
μM peptide instead of phage. Bound bi tinylated peptid was
detected with streptavidin conjugated to alkaline phosphatase
5 (Sigma Chemical Co., St. Louis, MO). After one hour
incubation period at room temperature, the wells were washed,
and a solution of 3 mM p-nitrophenyl-phosphate (US
Biochemicals, Cleveland, OH) in 50 mM NaCO$_3$ (pH 9.8), and 50
mM MgCl$_2$ was added and color allowed to develop. Signals wer
10 read with an ELISA plate reader (Molecular Devices, Menlo
Park, CA) at 405 nm wavelength. Binding experiments were
performed in triplicate. The results are presented in
Figures 7 and 8.

15   6.8. Peptide Competition of GST-SH3 Affinity
            Precipitations of Cell Lysates
        Labeled proteins are prepared by incubating a
culture of HeLa cells overnight with ≥100 μCi/mL ^{35}S-
methionine. The cells are then washed and lysed with mild
20 detergent. This mixture of radioactive proteins is incubated
with Src-GST fusion protein that has been immobilized on
glutathione-linked Sepharose beads (Pharmacia, Piscataway,
NJ). After several hours of tumbling, the beads are pelleted
gently by low-speed centrifugation, and the supernatant is
25 discarded. The beads are then resuspended into a slurry in
PBS-0.1% Tween 20, pelleted, and washed several additional
times. Finally, a 2% SDS solution is added to the sample,
which is then boiled at 100 °C for 3 minutes. Afterward, the
sample is centrifuged, and the supernatant loaded on a 10%
30 polyacrylamide SDS gel for electrophoresis. After the
proteins have been resolved, the gel is fixed, dried down,
and exposed to X-ray film for autoradiography or phosphor
plates for scanning by a Molecular Dynamics PhosphorImager.
        The ability of Src SH3 to bind certain ^{35}S-labeled
35 proteins is examin d f r comp tability with exogenous
peptid s. Synthetic peptides corresponding to phage-
displayed inserts and motifs are add d at the time that the

- 47 -

SUBSTITUTE SHEET (RULE 26)

lysat is incubat d with th Src-GST fusion protein
imm bilized on glutathione-linked sepharos beads. The SH3
binding p ptid s bl ck binding of all or some of the labeled
proteins while negative control peptides (unrelated peptide
5 sequences) do not. The amount of competition is quantified
and correlated with the amount of added SH3-domain binding
peptides.

        Alternatively, NIH 3T3 cells were grown in Dulbecco's
Modified Eagle Medium (DME) + 10% fetal calf serum (FCS) + 80
10 μCi/mL Tran^{35}Slabel (ICN), washed with PBS, lysed in RIPA
buffer, and pelleted. Supernatant from 1.5 x 10^6 cells was
precleared with 100 μg glutathione-agarose-immobilized GST.
The supernatant was then incubated with 10 μg glutathione-
agarose-immobilized GST-SH3 fusion protein with or without
15 added test peptide in a final volume of 250 μL. Pelleted
beads were washed with 1 mL each of RIPA, RIPA + 1%
deoxycholate + 0.1% SDS, and PBS, resuspended in 50 μL
SDS-PAGE sample buffer, boiled, and subjected to SDS-PAGE
(7.5%). Labeled proteins were detected by phosphorimaging
20 (Molecular Dynamics). The results are presented in Figure 9.

6.9. Peptide Competition of GST-SH3 Affinity Precipit-
ations of PI-3' Kinase From Cell Lysates

        It is possible to follow the precipitation of PI-3'
25 Kinase by Src from cell lysates in the presence or absence of
SH3-binding peptides. HeLa cells are lysed with detergent
and the protein mixtures are incubated for several hours with
the Src-GST fusion protein immobilized on glutathione-linked
Sepharose beads. After several hours of tumbling, the beads
30 are pelleted gently by low-speed centrifugation and the
supernatant is discarded. The beads are then resuspended
into a slurry in PBS-0.1% Tween 20, pelleted, and washed
several additional times. Finally, an SDS solution is added
to the sample, which is then boiled at 100 °C for 3 minutes.
35 Subs quently, th sample is c ntrifug d, and th supernatant
is loaded on a 10% polyacrylamide SDS gel for
  l ctrophor sis. Aft r the prot ins have been resolved, the

- 48 -

SUBSTITUTE SHEET (RULE 26)

g l is blott d to nitr cellulose or nylon (i. ., western
blot). Th filter is th n probed with a PI-3' Kinase
antibody (m noclonal and polycl nal antibodies are available
from Upstate Biotechnology Incorporated, Lake Placid, NY) and

5 an enzyme-linked secondary antibody. The amount of PI-3'
Kinase is then quantitated.

The ability of Src SH3 to bind PI-3' Kinase is examin d
for competability with exogenous peptides. Synthetic
peptides corresponding to phage-displayed inserts and motifs

10 are added at the time that the lysate is incubated with th
Src-GST fusion protein that has been immobilized on
glutathione-linked sepharose beads. Ten-fold and one
hundred-fold molar excess of peptides are used relative to
SH3 proteins. The SH3 binding peptides block binding of th

15 PI-3' Kinase as detected on western blots while negative
control peptides (unrelated peptide sequences) do not. The
amount of competition is quantified and correlated with the
amount of added SH3-domain binding peptides.

20   6.10.   *In Vivo* **Association Of SH3-Binding Peptides
                   With SH3-Domains Of Proteins**

To demonstrate association of the SH3-binding
peptides with SH3-domains of proteins inside cells, the SH3-
binding peptides are tagged and localized in cells. For

25 example, Bar-Sagi et al., in Cell (1993) 74:83-91, have shown
that SH3-binding proteins localize to the cytoskeleton when
expressed in cells. Thus, the SH3 domain-binding peptides of
the present invention can serve as cellular targetting
signals (e.g., to the cytoskeleton). Accordingly, the

30 peptides are tagged with biotin and, subsequently, injected
into cells. Alternatively, one can transfect into cells a
recombinant plasmid that expresses a fusion protein
comprising of the SH3-binding peptide and the green
fluorescent protein (GFP, Chalfie et al., in Science (1994)

35 263:802-805). The location f th biotinylated peptide or
th GFP fusion protein is th n assay d with FITC-labeled
streptavidin in paraformaldehyde-fix d cells or by direct

- 49 -

fluor sc nce in living cells, respectively. Localization of
the SH3-binding peptides to the cytoskeleton d monstrat s
that the SH3-binding peptides can bind SH3-domain proteins *in
vivo*. In addition, focal adhesions, which are rich in Src,
5 are also sites of potential subcellular localization of SH3-
binding peptides.

Thus, NIH 3T3 fibroblasts were cultured *in vitro* on
glass coverslips coated with fibronectin. After two days of
growth at 37 °C, the cells were fixed for one hour at room
10 temperature in the presence of 2% paraformaldehyde (pH 7.5).
The coverslips were washed with PBS-0.1% Tween 20 several
times to remove the fixative. Next, the coverslips were
dipped into acetone (chilled at -20 °C) for approximately 20
seconds and allowed to air-dry. The coverslips were washed
15 again with PBS-0.1% Tween 20, containing BSA (1 mg/mL), and
incubated for 2 hours at room temperature with different
biotinylated peptides in PBS-0.1% Tween 20. The coverslips
were washed and then incubated with 1 mg/mL streptavidin-Cy3
(Jackson Immunoresearch Co., West Grove, PA) for 1 hour at
20 room temperature. Finally, the coverslips were washed in
PBS-0.1% Tween 20, mounted in a glycerol solution on a glass
slide, and viewed with a Nikon Optiphot epifluorescence
microscope and a 60x oil immersion lens.

The results are presented in Figure 11, in which panel A
25 displays cells stained with the conjugate biotin-spacer-
VLKRPLPIPPVTR (SEQ ID NO:64); panel B exhibits cells stained
with the conjugate, biotin-spacer-GILAPPVPPRNTR (SEQ ID
NO:63); panel C shows cells stained with the long consensus
peptide, biotin-spacer-RSTPRPLPPLPTTR (SEQ ID NO:67); and
30 panel D shows cells stained with the proline-rich vinculin
peptide conjugate, biotin-spacer-LAPPKPPLPEGEV (SEQ ID
NO:70). The "spacer" sequence is KSGSG (SEQ ID NO:103). As
shown in Figure 11, the panels in which SH3 domain-binding
peptides were used present a bright display of fluorescence
35 activity that is in sharp contrast to th relatively "dark"
f atures of panel D (non-SH3 domain binding vinculin
s gment). Th s results demonstrate further the ability of

- 50 -

SUBSTITUTE SHEET (RULE 26)

the SH3 d main-binding peptides f the pres nt inv ntion to
localiz to pr tein targets (.g., Src and Src-relat d
proteins) within cells and provid an imag th reof.

5   6.11.   *In Vivo Modulation Of Src In Oocytes With*
                    SH3-Binding Peptides

                    When *Xenopus laevis* oocytes are injected with
mRNA encoding deregulated Src, there are *dramatic* cytological
and biochemical changes in the oocyte (Unger, T.F. and
10 Steele, R.E., in Mol. Cell. Biol. (1992) 12:5485-5498). The
applicants have obtained the plasmid for generating Src mRNA,
which is available from Dr. Robert Steele (University of
California at Irvine). Synthetic SH3-binding peptides are
injected into oocytes that have been previously injected with
15 Src mRNA. The state of the cytoskeleton is inspected
visually by observing the arrangement of cortical pigment
granules under a dissecting microscope. The state of
phosphorylation of several proteins is examined by western
blotting with an anti-phosphotryosine monoclonal antibody
20 (4G10; Upstate Biotechnology Incorporated), as described in
Unger and Steele, above.

        6.12.   **Progesterone-induced** *X. laevis* **Oocyte**
                    Maturation

                    Segments of adult ovary were removed
25 surgically and incubated in 0.1% collagenase type D
(Boehringer Mannheim, Indianapolis, IN) in Ca^{2+}-free OR2 (82.5
mM NaCl, 2.5 mM KCl, 1.0 mM $MgCl_2$, 1.0 mM Na_2HPO_4, 5.0 mM
HEPES, and 3.8 mM NaOH, pH 7.6). Oocytes were then washed 3-
5 times with OR2 containing 1.0 mM $CaCl_2$ and allowed to
30 recover in OR2 overnight at 18 °C. Stage VI oocytes were
injected with 40 nL of 100 mM peptide or water. After
injection, the oocytes were placed in OR2 with 2 mg/mL
progesterone (Sigma, St Louis, MO) and incubated at 20 °C.
Oocytes were scored at hourly time points for germinal
35 vesicle br akdown (GVBD).

- 51 -

Figure 10 pr sents th results of this experiment. As sh wn by th graph, oocyt s injected with th SH3 domain-binding peptid VLKRPLPIPPVTR (SEQ ID NO:64) exhibit a faster rate of progesterone-induced germinal vesicle breakdown

5 relative to oocytes that had been injected with water or with the proline-rich vinculin peptide, LAPPKPPLPEGEV (SEQ ID NO:70). These results parallel those obtained by Unger and Steele, *supra*, who observed a gross alteration in the cortex of oocytes that had been injected with RNA encoding a

10 deregulated Src variant ("active" Src) versus those injected with RNA encoding the wild-type Src ("cellular" Src). Also, as shown in Figure 3B of the Unger and Steele article, oocytes injected with deregulated or active Src RNA matured at a faster rate than oocytes injected with water or wild-

15 type Src RNA.

The present results obtained with Src SH3 domain-binding peptides suggest that these peptides modulate the biochemical activity of "cellular" Src; in particular, it is proposed that at least some of the Src SH3 domain-binding peptides of

20 the present invention upregulate the biochemical activity of "cellular" Src, which may be downregulated or inhibited in its normal state. Hence, the administration of the SH3 domain-binding peptides of the present invention can constitute a novel method of modulating the activity of Src

25 or Src-related proteins. Specifically, certain of these peptides are able to activate Src-family proteins.

6.13. *In Vivo* Antagonism Of Src In Src Transformed Cells With SH3-Binding Peptides

30 The coding regions for SH3-binding peptides are cloned into vectors that direct their expression in animal cells. A bipartite gene is constructed, encoding a protein with *c-myc* epitope and SH3-binding peptide, which is transcribed from a strong constitutive promoter (e.g., SV40,

35 CMV, HSV TK, calmodulin). The vector is introduced into either normal or Src-transformed c lls via transf ction (e.g., electroporation, calcium phosphate, liposomes, DEAE

- 52 -

SUBSTITUTE SHEET (RULE 26)

d xtran). Transf cted c lls expr ss th bipartite gene
transi ntly in cultur . To cr ate stabl transform d cell
lines, the vector carries a sel ctable marker (.g., n omycin
resistance) or transfection is performed in the presence of
5 excess plasmid carrying a selectable marker (e.g., neomycin
resistance) and cells selected for the marker. Transfected
cells are stained by immunofluorescence to detect expressi n
of the bipartite protein. The hybridoma 9E10 secretes a
monoclonal antibody that is highly specific for the *c-myc*
10 epitope (EQKLISEEDLN [SEQ ID NO:105]; see, Evan, G.A. et al.,
in Mol. Cell. Biol. (1985) 5:3610-3616). This antibody is
used in immunofluorescence experiments to demonstrate that
the bipartite protein is expressed inside the cells, and in
some cases, localized to subcellular structures enriched in
15 SH3 domain bearing proteins.

There are several controls used in these experiments.
First, cells are transfected with vectors that do not have
the SH3-binding peptide coding region. Second, normal (non-
transformed) cells are transfected. Third, cells transformed
20 by oncogenes other than Src are used in the transfection
experiments. Fourth, cells are stained with other monoclonal
antibodies that do not recognize the *c-myc* epitope.

Transfected cells are examined for any changes in cell
shape, behavior, and metabolism as a consequence of
25 expressing the SH3 binding peptides. Cell shape is examined
by phase contrast microscope at several times after
transfection; in particular, the flatness of the cells, their
adhesion to the substrate, and the degree of cell ruffling
are monitored. Cell division rates, cell migration, and
30 contact inhibition are also observed over time. Finally, the
amount of phosphorylated tyrosine in transfected cells is
quantitated by phosphoaminoacid analysis and with an anti-
phosphotryosine monoclonal antibody (4G10; Upstate
Biotechnology Incorporated) in western blotting experiments.
35

- 53 -

SUBSTITUTE SHEET (RULE 26)

6.14 Preparati n f PXXP Biased Peptide Librari s

The preparation and characterization of preferr d phag -
displayed rand m peptide libraries have be n described above
in Sections 6.1 - 6.4.

5   Using procedures similar to those described in these
sections, oligonucleotide inserts were contructed according
to the schematic provided in FIG. 12. The inserts were then
cloned into the mBAX vector, and the biased peptide libraries
were expressed as described previously.

10   The mBAX vector was created in the Kay Laboratory by
generating cloning sites in gene III of the M13mp18 vector
(Messing, J. (1991). Cloning in M13 phage or how to use
biology at its best. *Gene* 100, 3-12) in the manner of
Fowlkes et al. (1992)(*Biotechniques* 13, 422-427). The mBAX

15 vector displays a peptide sequence at the N-terminus of the
mature pIII protein that encodes the epitope for the mouse
monoclonal antibody 7E11 (see, FIG. 13); it includes the stop
codon TAG in the coding region, which is suppressed in *E.
coli* carrying suppressor tRNA gene mutations known as *supE* or

20 *supF*. There are no other stop codons in the mBAX genome.
The mBAX vector also carries a segment of the alpha fragment
of ß-galactosidase; bacterial cells expressing the omega
fragment of ß-galactosidase convert the clear XGal substrate
into an insoluble blue precipitate. The plaques appear blu .

25   Recombinant mBAX molecules can be distinguished easily
from non-recombinant molecules due to the TAG codon in the
*Xho*I - *Xba*I segment in gene III of mBAX. When recombinants
are generated by replacing the *Xho* I - *Xba* I fragment with
oligonucleotides encoding random peptides, the recombinants

30 can be grown in bacteria with (i.e., DH5αF') or without
(i.e., JS5) suppressor tRNA mutant genes. On the other hand,
the non-recombinant mBAX molecules fail to produce plaques on
bacterial lawns where the bacteria (i.e., JS5) lack such
suppressor genes. This is because in JS5, the TAG codon

35 serv s as a stop codon to yi ld a truncat d pIII molecule
during translation; since pIII is an essential prot in

- 54 -

SUBSTITUTE SHEET (RULE 26)

c mponent of viabl M13 viral particles, no plaqu s will
f rm.

      6.14.1   **Discussi n**

    The use of second generation or biased peptide

5 libraries, which fix all or part of the RPLPPLP consensus
motif for SH3 binding peptides and randomize flanking
residues, has defined additional sequence residues exhibiting
selective SH3 binding.

    Tables 1-5, below, list the relevant amino acid

10 sequences obtained from the biased peptide library for each
set of SH3 domain binding peptides. The underscored amino
acid residues indicate the fixed positions. Also, indicated
for each set of new binders is a "consensus" sequence, which
seeks to include the additional features gleaned from the new

15 binding peptides. The symbol "φ" represents a hydrophobic
residue.

TABLE 1   **CORTACTIN SH3-BINDING PEPTIDES**

		SEQ. ID NO.
PXXP.CORT.M1/2/3.PP	SSLLGPPVPPKPQTLFSFSR	107
PXXP.CORT.M4.PP	SRLGEFSKPPIPQKPTWMSR	108
PXXP.CORT.N2.PP	SRTERPPLPQRPDWLSYSSR	109
PXXP.CORT.N3.PP.INC	SREPDWLCPNCPLLLRSDSR	110
PXXP.CORT.01/2/3.PP	SSSSHNSRPPLPEKPSWLSR	111
PXXP.CORT.04.PP	SRLTPQSKPPLPPKPSAVSR	112
CONSENSUS	KPPφPxKPxW R	113

TABLE 2   NCK SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.NCK.Q1/4.PP	SSLGVGWKPLPPMRTASLSR	114
PXXP.NCK.Q2/3.PP.INC	SSVGFADRPRPPLRVESLSR	115
PXXP.NCK.R1.PP.INC	SSAGILRPPEKPXRSFSLSR	116
PXXP.NCK.R2.PP	SSPYTGDVPIPPLRGASLSR	117
PXXP.NCK.R3.PP	SSLMGSWPPVPPLRSDSLSR	118
PXXP.NCK.R4.PP	SSIGEDTPPSPPTRRASLSR	119
PXXP.NCK.S1/4.PP	SRSLSEVSPKPPIRSVSLSR	120
PXXP.NCK.S2.PP.INC	SSVSEGYSPPLPPRSTSLSR	121
PXXP.NCK.S3.PP	SSSFTLAAPTPPTRSLSLSR	122
PXXP.NCK.T1.PP	SSPPYELPPRPPNRTVSLSR	123
PXXP.NCK.T2.PP	SRVVDGLAPPPPVRLSSLSR	124
PXXP.NCK.T3.PP.INC	SSLGYSGAPVPPHRxSSLSR	125
PXXP.NCK.T4.PP	SSISDYSRPPPPVRTLSLSR	126
CONSENSUS	φXXXXXPxPPφRSxSL T	127

SUBSTITUTE SHEET (RULE 26)

TABLE 3   ABL SH3 BINDING PEPTIDES

		SEQ. ID NO.
PXXP.ABL.G1/2.PP	SRGPRWSPPPVPLPTSLDSR	128
PXXP.ABL.G3/4.PP	SSPPDYAAPAIPSSLWVDSR	129
PXXP.ABL.H1/3/4.PP	SSPPHWAPPAPPAMSPPISR	130
PXXP.ABL.H2.PP.INC	SSDRCWECPPWPAGGQRGSR	131
PXXP.ABL.I1/2/3.PP	SSPPKFSPPPPPYWQLHASR	132
PXXP.ABL.I4.PP	SSPPSFAPPAAPPRHSFGSR	133
PXXP.ABL.J1.PP	SSAPKKPAPPVPMMAHVMSR	134
PXXP.ABL.J2.PP.INC	SSPTYPPPPPPDTAKGASR	135
PXXP.ABL.J3.PP.INC	SSPPXXXPPPIPNSPQVLSR	136
PXXP.ABL.J4.PP	SSPPTWTPPKPPGWGVVFSR	137
PXXP.ABL.L1.PP	SSAPTWSPPALPNVAKYKSR	138
PXXP.ABL.L2/3.PP	SSIKGPRFPVPPVPLNGVSR	139
PXXP.ABL.L4.PP	SSPPAWSPPHRPVAFGSTSR	140
CONSENSUS	PPxWxPPPφP	141

TABLE 4   PLCG SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.PLCG.P1.PP	SSMKVHNFPLPPLPSYETSR	142
PXXP.PLCG.P2.PP	SRVPPLVAPRPPSTLNSLSR	143
PXXP.PLCG.PE.PP.INC	SSLYWQHGPDPPVGAPQLSR	144
PXXP.PLCG.P4.PP	SSHPLNSWPGGPFRHNLSSR	145

- 57 -

SUBSTITUTE SHEET (RULE 26)

TABLE 5   SRC SH3-BINDING PEPTIDES

		SEQ. ID NO.
PXXP.SRC.A1.PP	SSRALRVRPLPPVPGTSLSR	146
PXXP.SRC.A2.PP	SSFRALPLPPTPDNPFAGSR	147
PXXP.SRC.A3.PP	SRDAPGSLPFRPLPPVPTSR	148
PXXP.SRC.A4.PP	SSISQRALPPLPLMSDPASR	149
PXXP.SRC.B1.PP	SSPAYRPLPRLPDLSVIYSR	150
PXXP.SRC.B2/3/PP	SSFINRRLPALPPDNSLLSR	151
PXXP.SRC.B4.PP	SRLTGRPLPALPPPFSDFSR	152
PXXP.SRC.C1.PP	SRMKDRVLPPIPTVESAVSR	153
PXXP.SRC.C2.PP.INC	SSLYSAIAPDPPPRNSSSSR	154
PXXP.SRC.C3.PP	SSLASRPLPLLPNSAPGQSR	155
PXXP.SRC.D1.PP	SSLTSRPLPDIPVRPSKSSR	156
PXXP.SRC.D2.PP.INC	SSLKWRALPPLPETDTPYSR	157
PXXP.SRC.D3.PP	SSNTNRLPPPTPDGLDVRSR	158
PXXP.SRC.D4.PP	SSLQSRPLPLPPQSSYPISR	159
CONSENSUS	RPLPPLP	160

It should be apparent to one of ordinary skill that many other embodiments of the present invention can be contemplated beyond the preferred embodiments described above but which other embodiments nevertheless fall within the scope and spirit of the present invention. Hence, the present invention should not be construed to be limited to the preferred embodiments described herein, which serve only to illustrate the present invention, but only by the claims that follow.

Also, numerous references are cited throughout the specification. The complete disclosures of these references are incorporated by reference herein.

- 58 -

BNSDOCID: <WO___9603649A1_IB>

SEQUENCE LISTING

(1) ENERAL INFORMATION:

    (i) APPLICANT: SPARKS, ANDREW B.
                  KAY, BRIAN K.
                  THORN, JUDITH M.
                  QUILLIAM, LAWRENCE A.
                  DER, CHANNING J.

    (ii) TITLE OF INVENTION: Src SH3 BINDING PEPTIDES AND METHODS OF
         ISOLATING AND USING SAME

    (iii) NUMBER OF SEQUENCES: 106

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Pennie & Edmonds
        (B) STREET: 1155 Avenue of the Americas
        (C) CITY: New York
        (D) STATE: New York
        (E) COUNTRY: U.S.A.
        (F) ZIP: 10036-2711

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: S. Leslie Misrock
        (B) REGISTRATION NUMBER: 18,872
        (C) REFERENCE/DOCKET NUMBER: 1101-192

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (212) 790-9090
        (B) TELEFAX: (212) 869-9741/8864
        (C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 8
        (D) OTHER INFORMATION: /note= "hydrophobic residue (such
         as Pro or Leu)"

    (xi) SEQUENCE DESCRIPTI N: SEQ ID N :1:

Xaa Pr Xaa Xaa Pro Pro Pr Xaa Xaa Pro
1   5   10

- 59 -

SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:2:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 9 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

5   (ii) MOLECULE TYPE: peptide

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 1..3
          (D) OTHER INFORMATION: /note= "Xaa = any residue other
               than Cys"

10   (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 6
          (D) OTHER INFORMATION: /note= "X = any residue other
               than Cys"

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 8..9
          (D) OTHER INFORMATION: /note= "X = any residue other
15   than Cys"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

     Xaa Xaa Xaa Pro Pro Xaa Pro Xaa Xaa
     1   5

(2) INFORMATION FOR SEQ ID NO:3:

20   (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 9 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
25   (B) LOCATION: 2
          (D) OTHER INFORMATION: /note= "X tends to be hydrophobic"

     (ix) FEATURE:
          (A) NAME/KEY: Modified-site
          (B) LOCATION: 8..9
          (D) OTHER INFORMATION: /note= "X tends to be hydrophobic"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
30
     Arg Xaa Leu Pro Pro Arg Pro Xaa Xaa
     1   5

(2) INFORMATION FOR SEQ ID NO:4:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 9 amino acids
35   (B) TYPE: amin acid
          (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

- 60 -

SUBSTITUTE SHEET (RULE 26)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Lys Leu Pro Pr Arg Pro Arg Arg
1   5

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

    Val Pro Pro Pro Val Pro Pro Arg Arg Arg
    1   5   10

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "tends to be Pro"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /note= "hydrophobic residue"

   (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "tends to be Pro"

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

    Xaa Xaa Xaa Pro Xaa Xaa Pro
    1   5

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (ix) FEATURE:
        (A) NAME/KEY: Modified-sit
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "X = any residue other
             than Cys"

- 61 -

SUBSTITUTE SHEET (RULE 26)

        (ix) FEATURE:
             (A) NAME/KEY: Modified-site
             (B) LOCATION: 8..9
             (D) OTHER INFORMATION: /n t = "X = any r sidue other
                     than Cys"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

5

     Arg Xaa Leu Pro Pro Arg Pro Xaa Xaa
     1   5

     (2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 9 amino acids
10   (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (ix) FEATURE:
             (A) NAME/KEY: Modified-site
             (B) LOCATION: 9
             (D) OTHER INFORMATION: /note= "hydrophobic residue"

15
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

     Arg Xaa Leu Pro Pro Leu Pro Arg Phe
     1   5

     (2) INFORMATION FOR SEQ ID NO:9:

20   (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 7 amino acids
             (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

25   Arg Pro Leu Pro Pro Leu Pro
     1   5

     (2) INFORMATION FOR SEQ ID NO:10:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 9 amino acids
             (B) TYPE: amino acid
30   (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (ix) FEATURE:
             (A) NAME/KEY: Modified-site
             (B) LOCATION: 2
             (D) OTHER INFORMATION: /note= "any residue other than Cys"

35   (ix) FEATURE:
             (A) NAME/KEY: Modified-site
             (B) LOCATION: 5..6
             (D) OTHER INFORMATION: /n t = "any hydr phobic residue"

                              - 62 -

                    SUBSTITUTE SHEET (RULE 26)

```
        (ix) FEATURE:
              (A) NAME/KEY: Modified-site
              (B) LOCATION: 9
              (D) OTHER INFORMATION: /not = "any hydrophilic residu
                      except Cys"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
5
        Arg Xaa Leu Pro Xaa Xaa Pro Xaa Xaa
        1               5


   (2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 9 amino acids
10            (B) TYPE: amino acid
              (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

        Arg Pro Leu Pro Pro Leu Pro Thr Ser
        1               5
15
   (2) INFORMATION FOR SEQ ID NO:12:

        (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 4 amino acids
              (B) TYPE: amino acid
              (D) TOPOLOGY: linear

20      (ii) MOLECULE TYPE: peptide

        (v) FRAGMENT TYPE:

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

        Xaa Xaa Xaa Xaa
        1


25
   (2) INFORMATION FOR SEQ ID NO:13:

        (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 5 amino acids
              (B) TYPE: amino acid
              (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide
30
        (v) FRAGMENT TYPE:

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

        Xaa Xaa Xaa Xaa Xaa
        1               5


35 (2) INFORMATION FOR SEQ ID NO:14:

        (i) SEQUENCE CHARACTERISTICS:
              (A) LENGTH: 4 amin  acids
              (B) TYPE: amino acid
```

- 63 -

SUBSTITUTE SHEET (RULE 26)

                    (D) TOPOLOGY: linear

              (ii) MOLECULE TYPE: peptide

               (v) FRAGMENT TYPE:

              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

       Xaa Xaa Xaa Xaa
       1

        (2) INFORMATION FOR SEQ ID NO:15:

              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 13 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: unknown

              (ii) MOLECULE TYPE: peptide

              (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 1..3
                    (D) OTHER INFORMATION: /note= "any residue other than Cys"

              (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 5
                    (D) OTHER INFORMATION: /note= "any residue other than Cys"

              (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 8..9
                    (D) OTHER INFORMATION: /note= "any hydrophobic residue"

              (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 11..13
                    (D) OTHER INFORMATION: /note= "any residue other than Cys"

              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

       Xaa Xaa Xaa Arg Xaa Leu Pro Xaa Xaa Pro Xaa Xaa Xaa
       1   5   10

        (2) INFORMATION FOR SEQ ID NO:16:

              (i) SEQUENCE CHARACTERISTICS:
                    (A) LENGTH: 42 amino acids
                    (B) TYPE: amino acid
                    (D) TOPOLOGY: linear

              (ii) MOLECULE TYPE: peptide

              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

       Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
       1   5   10   15

       Xaa Xaa Xaa Xaa Pr Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                       20   25   30

       Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg

                            - 64 -

35   40

(2) INFORMATI N FOR SEQ ID NO:17:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 69 base pairs
            (B) TYPE: nucleic acid
  5   (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

    GGCTCGAGNN NBNNBNNBNN BNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNNBNNBNNB   60
10 NNBCCAGGT   69

(2) INFORMATION FOR SEQ ID NO:18:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 68 base pairs
            (B) TYPE: nucleic acid
 15   (C) STRANDEDNESS: both
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

    GGTCCANNVN NVNNVNNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVNN VNNVNNVNNV   60
 20 AGATCTGG   68

(2) INFORMATION FOR SEQ ID NO:19:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 66 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
 25   (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

    TCGAGNNNBN NBNNBNNBNN BNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNNBNNBNNB   60
    CCAGGT   66
 30

(2) INFORMATION FOR SEQ ID NO:20:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 65 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

 35
        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

- 65 -

SUBSTITUTE SHEET (RULE 26)

```
GGTCCANNVN NVNNVNNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVNN VNNVNNVNNV     60

AGATC                                                                65
```

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 46 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (v) FRAGMENT TYPE:

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "Arg or Ser"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1               5                   10                  15

Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Pro Ser Arg Thr
            35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 49 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Ser His Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1               5                   10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            20                  25                  30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Arg Pro Ser Arg
            35                  40                  45

Thr
```

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:

- 66 -

SUBSTITUTE SHEET (RULE 26)

                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 2
                    (D) OTHER INFORMATION: /note= "Ser or Thr"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 13
     5   (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

            (ix) FEATURE:
                    (A) NAME/KEY: Modified-site
                    (B) LOCATION: 15
                    (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

    10   Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa
           1   5   10   15

           Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg
                         20   25

        (2) INFORMATION FOR SEQ ID NO:24:

            (i) SEQUENCE CHARACTERISTICS:
    15   (A) LENGTH: 48 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: both
                (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: DNA (genomic)

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

    20 TTTTGTCGAC NNNBNNBNNB NNBNNBNNBN NBNNBNNBNN BNGCGGTG   48

        (2) INFORMATION FOR SEQ ID NO:25:

            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 48 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: both
    25   (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: DNA (genomic)

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

        CGCCACNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVTG ATCATTTT   48

    30 (2) INFORMATION FOR SEQ ID NO:26:

            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 43 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: double
                (D) TOPOLOGY: linear

            (ii) MOLECULE TYPE: DNA (genomic)
    35
            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

        CGCCACNVNN VNNVNNVNNV NNVNNVNNVN NVNNVNNVTG ATC   43

                              - 67 -

                        SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 43 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

5

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCGACNNNBN NBNNBNNBNN BNNBNNBNNB NNBNNBNGCG GTG   43

(2) INFORMATION FOR SEQ ID NO:28:

10

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (v) FRAGMENT TYPE:

15

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "Ser or Thr"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 13

20

        (D) OTHER INFORMATION: /note= "Ser, Arg, Gly, Cys or Trp"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 15
        (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

25

Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa
1   5   10   15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg Pro Ser Arg Thr
       20   25   30

(2) INFORMATION FOR SEQ ID NO:29:

30

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 34 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site

35

        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "Ser or Thr"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site

- 68 -

SUBSTITUTE SHEET (RULE 26)

                              (B) LOCATION: 16
                              (D) OTHER INFORMATI N: /n t = "Ser, Arg, Gly, Cys or Trp"

                    (ix) FEATURE:
                              (A) NAME/KEY: Modified-site
                              (B) LOCATION: 18
                              (D) OTHER INFORMATION: /note= "Val, Ala, Asp, Glu or Gly"
5

                    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

                    Ser His Ser Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                    1   5   10   15

                    Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Arg Pro Ser
                              20   25   30
10
                    Arg Thr

          (2) INFORMATION FOR SEQ ID NO:30:

                    (i) SEQUENCE CHARACTERISTICS:
15   (A) LENGTH: 8 amino acids
                          (B) TYPE: amino acid
                          (D) TOPOLOGY: linear

                    (ii) MOLECULE TYPE: peptide

                    (v) FRAGMENT TYPE: internal

                    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
20
                    Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                    1   5

          (2) INFORMATION FOR SEQ ID NO:31:

                    (i) SEQUENCE CHARACTERISTICS:
                          (A) LENGTH: 58 base pairs
25   (B) TYPE: nucleic acid
                          (C) STRANDEDNESS: both
                          (D) TOPOLOGY: unknown

                    (ii) MOLECULE TYPE: DNA (genomic)

                    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

          TGACGTCTCG AGTTGTNNKN NKNNKNNKNN KNNKNNKNNK TGTGGATCTA GAAGGATC   58
30
          (2) INFORMATION FOR SEQ ID NO:32:

                    (i) SEQUENCE CHARACTERISTICS:
                          (A) LENGTH: 15 base pairs
                          (B) TYPE: nucleic acid
                          (C) STRANDEDNESS: both
                          (D) TOPOLOGY: unknown
35
                    (ii) MOLECULE TYPE: DNA (genomic)

                    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

                                        - 69 -

CCTAGATCTT CCTAG   15

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 40 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TCGAGTTGTN NKNNKNNKNN KNNKNNKNNK NNKTGTGGAT   40

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 40 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CAACANNMNN MNNMNNMNNM NNMNNMNNMA CACCTAGATC   40

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (v) FRAGMENT TYPE:

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

    Ser Ser Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Gly Ser Arg Pro
    1   5   10   15

    Ser Arg Thr

(2) INFORMATION FOR SEQ ID NO:36:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

    Ser Ser Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Gly Ser Arg Ser
    1   5   10   15

- 70 -

SUBSTITUTE SHEET (RULE 26)

```
        Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Thr Arg
                 20                      25
```

     (2) INFORMATION FOR SEQ ID NO:37:

         (i) SEQUENCE CHARACTERISTICS:
5   (A) LENGTH: 80 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: double
             (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
10
TCGAGTTGTN NKNNKNNKNN KNNKNNKNNK NNKTGTGGAT CTAGATCCAC AVNNVNNVNN   60

VNNVNNVNNV NNVNNACAAC   80

     (2) INFORMATION FOR SEQ ID NO:38:

         (i) SEQUENCE CHARACTERISTICS:
15   (A) LENGTH: 45 amino acids
             (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
        Ser Ser Phe Asp Gln Gln Asp Trp Asp Tyr Ser Ile Ala Glu Lys Met
20      1               5                  10                  15

        His Pro Ile Arg Pro Gly Phe Arg Glu Leu Pro Pro Leu Pro Pro Ser
                     20                  25                  30

        Arg Ala Ser Phe Gly Gly Gly Ala Ser Arg Pro Ser Arg
                 35                  40                  45
```

25 (2) INFORMATION FOR SEQ ID NO:39:

         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 30 amino acids
             (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

30   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
        Ser Thr Asn Val Trp Val Thr Gly Ser Val Ile Ala Arg Gly Ala Gln
        1               5                  10                  15

        Ser Arg Pro Leu Pro Ile Pro Pro Glu Thr Arg Pro Ser Arg
                     20                  25                  30
```

35 (2) INFORMATION FOR SEQ ID NO:40:

         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 30 amino acids
             (B) TYPE: amin acid

                              - 71 -

                    SUBSTITUTE SHEET (RULE 26)

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ser Thr Ala Pro Trp Gly Leu Arg Val Ala His Glu Gly Gly Val Leu
1   5   10   15

Lys Arg Pro Leu Pro Ile Pro Pro Val Thr Arg Pro Ser Arg
            20   25   30

(2) INFORMATION FOR SEQ ID NO:41:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 45 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ser Ser Ser Gly Tyr Val Val Pro Lys Arg Leu Gly Asp Met Arg Glu
1   5   10   15

Tyr Asn Ala His Pro Gly Leu His Val Pro Pro Asn Ser Pro Leu Pro
            20   25   30

Pro Leu Pro Thr His Leu Gln Ser Ser Arg Pro Ser Arg
        35   40   45

(2) INFORMATION FOR SEQ ID NO:42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 44 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Ser Ser Arg Gly Glu Gly Asn Asn Ile Ile Ser Ser Arg Pro Phe Leu
1   5   10   15

Ser Asn Ser Asp Pro Gly Val Ser Asn Lys Leu Thr Gly Arg Gly Pro
            20   25   30

Leu Pro Pro Leu Pro Asn Asp Ser Arg Pro Ser Arg
        35   40

(2) INFORMATION FOR SEQ ID NO:43:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

- 72 -

SUBSTITUTE SHEET (RULE 26)

```
Ser Thr Ala Val S r Ph  Arg Phe Met Pr  Gly Gly Gly Gly Ala Ph
1              5                   10               15

Tyr Ser Thr Arg Pro Val Pro Pr  Ile Thr Arg Pro Ser Arg Thr
            20              25               30
```

(2) INFORMATION FOR SEQ ID NO:44:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```
Ser Thr Ala His Ser Leu Trp Asp Trp Gly Thr Phe Ser Gly Val Ser
1              5                   10               15

His Lys Ser Arg Leu Pro Pro Leu Pro Thr Arg Pro Ser Arg Thr
            20              25               30
```

(2) INFORMATION FOR SEQ ID NO:45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Pro Gly Tyr Ala Arg Ile Val Ser Tyr Arg Phe Arg Ala Leu Pro Ser
1              5                   10               15

Pro Pro Ser Ala Ser Arg Pro Ser Arg
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
Ser Thr Asn Asp Val Asp Trp Met His Met Trp Asn Ser Gly Gly Pro
1              5                   10               15

His Arg Arg Leu Pro Pro Thr Pro Ala Thr Arg Pro Ser Arg
            20              25               30
```

(2) INFORMATION FOR SEQ ID NO:47:

- 73 -

SUBSTITUTE SHEET (RULE 26)

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 45 amin acids
    (B) TYPE: amin acid
    (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

5   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Ser Ser Asp Asn Trp Ala Arg Arg Val His Ala Ser Glu Leu Ile Tyr
1   5   10   15

Thr Asp Leu Ser Pro Gly Ile Leu Leu Ala Gln Arg Gln Leu Pro Pro
           20   25   30

Thr Pro Gly Arg Asp Pro Ser His Ser Arg Pro Ser Arg
        35   40   45

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 45 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Ser Ser Glu Ser Pro Leu Met Tyr Asn Arg Val Gly Ala Leu Gln Ser
1   5   10   15

Leu Thr Ser Val Pro Gly Ser Met Met His Phe Ala Leu Gln Arg Arg
           20   25   30

Leu Pro Arg Thr Pro Pro Pro Ala Ser Arg Pro Ser Arg
        35   40   45

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 35 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Ser Thr Arg Trp Ser His Ser Trp Pro Gly Tyr Val Gly Gly Ala Asn
1   5   10   15

Pro Ser Pro Ala Thr Arg Pro Leu Pro Thr Arg Pro Ser Arg Thr Val
           20   25   30

Glu Ser Cys
    35

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 50 amin acids
    (B) TYPE: amino acid

- 74 -

SUBSTITUTE SHEET (RULE 26)

```
            (D) TOPOLOGY: unknown

     (ii) MOLECULE TYPE: peptide

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

     Ser Arg Tyr Asn Asp Leu Gly Thr Arg Pro Val Ser Glu Val Ile Lys
     1               5                   10                  15

     Tyr Asp Tyr Phe Pro Gly Tyr Ser Gln His Val Ile Thr Pro Asp Gly
                 20                  25                  30

     Ser Tyr Ser Thr Arg Pro Leu Pro Ser Arg Pro Ser Arg Thr Val Glu
             35                  40                  45

     Ser Cys
         50


     (2) INFORMATION FOR SEQ ID NO:51:

         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 24 amino acids
             (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

         (ii) MOLECULE TYPE: peptide

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

         Pro Gly Arg Leu Leu Leu Pro Ser Glu Pro Arg Thr Phe Tyr Asn Tyr
         1               5                   10                  15

         Gly His Asp Ser Arg Pro Ser Arg
                         20


     (2) INFORMATION FOR SEQ ID NO:52:

         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 30 amino acids
             (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

         (ii) MOLECULE TYPE: peptide

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

         Ser Thr Met Tyr Gly Val Ser Trp Leu Ser Ser Gly Ser Gly Gly Ile
         1               5                   10                  15

         Leu Ala Pro Pro Val Pro Pro Arg Asn Thr Arg Pro Ser Arg
                     20                  25                  30


     (2) INFORMATION FOR SEQ ID NO:53:

         (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 31 amino acids
             (B) TYPE: amino acid
             (D) TOPOLOGY: unknown

         (ii) MOLECULE TYPE: peptide

         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:
```

- 75 -

```
Ser Ser Cys Thr Glu Lys Thr Val Ser Gly Trp Cys Gly Ser Arg Ser
1             5                 10               15

Thr Arg Pr  Leu Pr  Ile Leu Pr  Arg Thr Thr Arg Pro Ser Arg
            20              25              30
```

5 (2) INFORMATION FOR SEQ ID NO:54:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 31 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

       (ii) MOLECULE TYPE: peptide

10   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```
Ser Ser Cys Met Leu Pro Thr Asp Gly Trp Gln Cys Gly Ser Arg Ser
1             5                 10               15

Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg Pro Ser Arg
            20              25              30
```

15 (2) INFORMATION FOR SEQ ID NO:55:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 31 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

       (ii) MOLECULE TYPE: peptide

20   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```
Ser Ser Cys Asp Gly Thr Gln Phe Arg Leu Asn Cys Gly Ser Arg Ser
1             5                 10               15

Thr Asn Arg Pro Leu Pro Met Ile Pro Thr Thr Arg Pro Ser Arg
            20              25              30
```

25 (2) INFORMATION FOR SEQ ID NO:56:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 31 amino acids
           (B) TYPE: amino acid
           (D) TOPOLOGY: unknown

       (ii) MOLECULE TYPE: peptide

30   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```
Ser Ser Cys Met Gln Gly Gln Ala Gly Leu Lys Cys Gly Ser Arg Ser
1             5                 10               15

Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr Arg Pro Ser Arg
            20              25              30
```

35

    (2) INFORMATION FOR SEQ ID NO:57:

- 76 -

    (i) SEQUENCE CHARACTERISTICS:
        (A) LEN TH: 31 amin acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

5   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Ser Ser Cys Tyr Arg Glu Lys Asp Thr Trp Gly Cys Gly Ser Arg Ser
1   5   10   15

Thr Ser Arg Pro Leu Pro Ser Leu Pro Thr Thr Arg Pro Ser Arg
        20   25   30

10 (2) INFORMATION FOR SEQ ID NO:58:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 32 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

15   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Ser Ser Cys Leu Phe Glu Gln Gly Ala Gly Thr Cys Gly Ser Arg Ser
1   5   10   15

Thr Arg Ser Leu Pro Pro Leu Pro Pro Thr Thr Arg Pro Ser Ser Arg
        20   25   30

20 (2) INFORMATION FOR SEQ ID NO:59:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 44 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

25   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Ser Ser Cys Asp Thr Gly Arg Ile Ala Pro Gly Cys Gly Ser Arg Ser
1   5   10   15

Thr Pro Arg Pro Leu Pro Leu Ile Pro Thr Thr Pro Arg Ser Thr Asn
        20   25   30

Leu Asn Leu Thr Ser Thr Thr Thr Arg Pro Ser Arg
        35   40
30

    (2) INFORMATION FOR SEQ ID NO:60:

        (i) SEQUENCE CHARACTERISTICS:
35   (A) LENGTH: 31 amin acids
         (B) TYPE: amino acid
         (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptid

- 77 -

SUBSTITUTE SHEET (RULE 26)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
Ser Ser Cys Gly Leu Asp Asn Ala Ala Lys Thr Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Pro Leu Pro Pro Thr Pro Leu Thr Thr Arg Pro Ser Arg
            20                  25                  30
```

5

   (2) INFORMATION FOR SEQ ID NO:61:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 31 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

10   (ii) MOLECULE TYPE: peptide

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Ser Ser Cys Ser Arg Ala His Glu Thr Glu Met Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Pro Gln Pro Pro Pro Pro Ile Thr Thr Arg Pro Ser Arg
            20                  25                  30
```

15

   (2) INFORMATION FOR SEQ ID NO:62:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 14 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

20   (ii) MOLECULE TYPE: peptide

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```
Arg Ser Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg
1               5                   10
```

   (2) INFORMATION FOR SEQ ID NO:63:
25
      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 13 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: peptide

30
      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```
Gly Ile Leu Ala Pro Pro Val Pro Pro Arg Asn Thr Arg
1               5                   10
```

   (2) INFORMATION FOR SEQ ID NO:64:

35   (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 13 amin acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: unknown

- 78 -

SUBSTITUTE SHEET (RULE 26)

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Val Leu Lys Arg Pro Leu Pro Ile Pro Pro Val Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:65:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Gly Pro His Arg Arg Leu Pro Pro Thr Pro Ala Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:66:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Ala Asn Pro Ser Pro Ala Thr Arg Pro Leu Pro Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:67:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Arg Ser Thr Pro Arg Pro Leu Pro Pro Leu Pro Thr Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:68:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

- 79 -

SUBSTITUTE SHEET (RULE 26)

```
Gly Thr Val  lu Pr  Val Pro Pro Pro Val Pr  Pro Arg Arg Arg Pr
1                5                    10           15

 lu Ser Ala
```

(2) INFORMATION FOR SEQ ID NO:69:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```
Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro Arg Pro Leu Pro
1                5                    10           15

Val
```

(2) INFORMATION FOR SEQ ID NO:70:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```
Leu Ala Pro Pro Lys Pro Pro Leu Pro Glu Gly Glu Val
1                5                    10
```

(2) INFORMATION FOR SEQ ID NO:71:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 2
        (D) OTHER INFORMATION: /note= "any residue other than Cys"

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5..6
        (D) OTHER INFORMATION: /note= "any hydrophobic residue"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```
Arg Xaa Leu Pro Xaa Xaa Pro
1                5
```

(2) INFORMATION FOR SEQ ID NO:72:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids

- 80 -

SUBSTITUTE SHEET (RULE 26)

(B) TYPE: amin acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Pro Gly Phe Arg Glu Leu Pro Pro Leu Pro Pro Ser Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Ala Gln Ser Arg Pro Leu Pro Ile Pro Pro Glu Thr Arg
1   5   10

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Pro Pro Asn Ser Pro Leu Pro Pro Leu Pro Thr His Leu
1   5   10

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Thr Gly Arg Gly Pro Leu Pro Pro Leu Pro Asn Asp Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

- 81 -

SUBSTITUTE SHEET (RULE 26)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Tyr Ser Thr Arg Pro Val Pr Pro Ile Thr Arg Pr Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:77:
5
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:
10
        Ser His Lys Ser Arg Leu Pro Pro Leu Pro Thr Arg Pro
        1   5   10

(2) INFORMATION FOR SEQ ID NO:78:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
15   (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

        Tyr Arg Phe Arg Ala Leu Pro Ser Pro Pro Ser Ala Ser
        1   5   10
20
(2) INFORMATION FOR SEQ ID NO:79:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

25   (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

        Leu Ala Gln Arg Gln Leu Pro Pro Thr Pro Gly Arg Asp
        1   5   10

(2) INFORMATION FOR SEQ ID NO:80:
30
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:
35
        Ala Leu Gln Arg Arg Leu Pro Arg Thr Pr Pr Pro Ala
        1   5   10

- 82 -

SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:81:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amin acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

5   (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

    Pro Ala Thr Arg Pro Leu Pro Thr Arg Pro Ser Arg Thr
    1   5   10

(2) INFORMATION FOR SEQ ID NO:82:

10   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

15   Tyr Ser Thr Arg Pro Leu Pro Ser Arg Pro Ser Arg Thr
    1   5   10

(2) INFORMATION FOR SEQ ID NO:83:

    (i) SEQUENCE CHARACTERISTICS:
20   (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

    Xaa Pro Gly Arg Ile Leu Leu Leu Pro Ser Glu Pro Arg
25   1   5   10

(2) INFORMATION FOR SEQ ID NO:84:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

30   (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

    Ser Gly Gly Ile Leu Ala Pro Pro Val Pro Pro Arg Asn
    1   5   10

35 (2) INFORMATION FOR SEQ ID NO:85:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amin acids
        (B) TYPE: amin acid

- 83 -

SUBSTITUTE SHEET (RULE 26)

              (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: peptid

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

      Arg Ser Thr Arg Pro Leu Pro Ile Leu Pro Arg Thr Thr
 5   1   5   10

    (2) INFORMATION FOR SEQ ID NO:86:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown
10
        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

        Ser Thr Pro Arg Pro Leu Pro Met Leu Pro Thr Thr Arg
        1   5   10

15

    (2) INFORMATION FOR SEQ ID NO:87:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

20   (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

        Ser Thr Asn Arg Pro Leu Pro Met Ile Pro Thr Thr Arg
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:88:
25
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:
30
        Arg Ser Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:89:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
35   (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptid

                          - 84 -

                   SUBSTITUTE SHEET (RULE 26)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

        Ser Thr Ser Arg Pro Leu Pr Ser Leu Pr Thr Thr Arg
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:90:
5
        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:
10
        Arg Ser Thr Arg Pro Leu Pro Ser Leu Pro Ile Thr Thr
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:91:

15   (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

20   Arg Ser Thr Arg Gln Leu Pro Ile Pro Pro Thr Thr Thr
        1   5   10

    (2) INFORMATION FOR SEQ ID NO:92:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
25   (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

        Ser Thr Pro Arg Pro Leu Pro Leu Ile Pro Thr Thr Pro
        1   5   10
30
    (2) INFORMATION FOR SEQ ID NO:93:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 13 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: unknown

35   (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

        Arg Ser Thr Arg Pro Leu Pr Pr Thr Pr Leu Thr Thr

                            - 85 -

SUBSTITUTE SHEET (RULE 26)

1   5   10

(2) INFORMATION FOR SEQ ID NO:94:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

    Arg Ser Thr Arg Pro Gln Pro Pro Pro Ile Thr Thr
    1   5   10

(2) INFORMATION FOR SEQ ID NO:95:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

    Arg Pro Leu Pro Met Leu Pro
    1   5

(2) INFORMATION FOR SEQ ID NO:96:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 7 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (ix) FEATURE:
        (A) NAME/KEY: Modified-site
        (B) LOCATION: 5
        (D) OTHER INFORMATION: /note= "Pro or Met"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

    Arg Pro Leu Pro Xaa Leu Pro
    1   5

(2) INFORMATION FOR SEQ ID NO:97:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

- 86 -

SUBSTITUTE SHEET (RULE 26)

Arg Ser Thr Pr
1

(2) INFORMATION FOR SEQ ID NO:98:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

    Arg Ser Thr Pro Ala Pro Pro Val Pro Pro Arg Thr Thr Arg
    1   5   10

(2) INFORMATION FOR SEQ ID NO:99:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

AGCGTAACGA TCTCCCG   17

(2) INFORMATION FOR SEQ ID NO:100:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TTCACCTCGA AAGCAAGCTG   20

(2) INFORMATION FOR SEQ ID NO:101:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CCTCATAGTT AGCGTAACG   19

- 87 -

SUBSTITUTE SHEET (RULE 26)

(2) INFORMATION FOR SEQ ID NO:102:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 base pairs
        (B) TYPE: nucl ic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: linear

5

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

AGCGTAACGA TCTAAA   16

10 (2) INFORMATION FOR SEQ ID NO:103:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

15   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

Lys Ser Gly Ser Gly
1   5

(2) INFORMATION FOR SEQ ID NO:104:

20   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

25 Gly Ser Gly Ser
1

(2) INFORMATION FOR SEQ ID NO:105:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 11 amino acids
        (B) TYPE: amino acid
30   (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
1   5   10
35

(2) INFORMATION FOR SEQ ID N :106:

    (i) SEQUENCE CHARACTERISTICS:

- 88 -

SUBSTITUTE SHEET (RULE 26)

(A) LENGTH: 31 amino acids
(B) TYPE: amin acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

```
Ser Ser Cys Asp His Thr Leu Gly Leu Gly Trp Cys Gly Ser Arg Ser
1               5                   10                  15

Thr Arg Gln Leu Pro Ile Pro Pro Thr Thr Thr Arg Pro Ser Arg
              20                  25                  30
```

- 89 -

SUBSTITUTE SHEET (RULE 26)

WHAT IS CLAIMED IS:

1.   A peptide having at least nine and up t f rty-five
amino acid residues, including an amin acid sequence of the
formula, R-2-L-P-5-6-P-8-9 (SEQ ID NO:10), positioned
5 anywhere along the peptide, in which each number represents
an amino acid residue, such that 2 represents any amino acid
residue except cysteine, 5 and 6 each represents a
hydrophobic amino acid residue, 8 represents any amino acid
residue except cysteine, and 9 represents a hydrophilic amino
10 acid residue except cysteine, each letter being the standard
one-letter symbol for the corresponding amino acid, said
peptide exhibiting a binding affinity for the SH3 domain of
Src, provided that said peptide is not R-P-L-P-P-L-P-T-S (SEQ
ID NO:11).

15   2.   The peptide of claim 1 in which 2 is a P, R, A, L,
Q, E or S.

3.   The peptide of claim 1 in which 5 is a P, M, I or
L.

4.   The peptide of claim 1 in which 6 is a P, L, I or
20 V.

5.   The peptide of claim 1 in which 8 is a T, R, P, I,
N, E, V, S, A, G or L.

6.   The peptide of claim 1 in which 9 is a T, R, S, H
or D.

25   7.   The peptide of claim 1 which further comprises a
C-terminal-flanking amino acid sequence of the formula 10,
10-11, 10-11-12, 10-11-12-13 (SEQ ID NO:12) or 10-11-12-13-14
(SEQ ID NO:13), in which each number represents any amino
acid residue except cysteine, such that 10 is bound to 9 by a
30 peptide bond.

8.   The peptide of claim 7 in which 10 is T, R, L, S,
D, P, A or N.

9.   The peptide of claim 7 in which 11 is R, P, A, Q, S
or T.

35   10.   The peptid of claim 7 in which 12 is P, S, R or T.

11.   The peptide of claim 7 in which 13 is P, S, R, F, H
or T.

- 90 -

SUBSTITUTE SHEET (RULE 26)

12. Th peptide of claim 7 in which 14 is S, R, G or T.

13. Th peptid   f claim 1 which furth r compris s an
N-terminal-flanking amin acid sequence f th formula 1',
2'-1', 3'-2'-1' or 4'-3'-2'-1' (SEQ ID NO:14) in which each
5 number represents any amino acid residue except cysteine,
such that 1' is bound to R by a peptide bond.

14. The peptide of claim 13 in which 1' is T, P, S, N,
F, W, K, H, Q or G.

15. The peptide of claim 13 in which 2' is S, T, G, P,
10 R, Q, L, A or H.

16. The peptide of claim 13 in which 3' is R, S, P, G,
A, V, Y or L.

17. The peptide of claim 13 in which 4' is R, S, V, T,
G, L or F.

15   18. A peptide having at least thirteen and up to forty-
five amino acid residues, including an amino acid sequence of
the formula, 3'-2'-1'-R-2-L-P-5-6-P-8-9-10 (SEQ ID NO:15),
positioned anywhere along the peptide, in which each number
represents an amino acid residue, such that 3', 2', 1', 2, 8,
20 and 10 each represents any amino acid residue except
cysteine, 5 and 6 each represents a hydrophobic amino acid
residue, and 9 represents a hydrophilic amino acid residue
except cysteine, each letter being the standard one-letter
symbol for the corresponding amino acid, said peptide
25 exhibiting a binding affinity for the SH3 domain of Src.

19. The peptide of claim 18 in which 5 is a P or M.

20. The peptide of claim 18 in which 1' is T, P, S or
N.

21. The peptide of claim 18 in which 2' is S or T.

30 22. The peptide of claim 18 in which 3' is R or S.

23. The peptide of claim 18 in which 10 is T or R.

24. The peptide of claim 1 the binding affinity of
which is at least three-fold greater than that exhibited by
the peptide RPLPPLP for the SH3 domain of Src.

35 25. The peptide f claim 18 th binding affinity f
which is at 1 ast three-f ld gr ater than that exhibit d by
the p ptide RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

- 91 -

SUBSTITUTE SHEET (RULE 26)

26. The p ptid f claim 1 th binding affinity f
which is at 1 ast four-fold gr ater than that xhibited by
the peptide RPLPPLP (SEQ ID NO:9) for th SH3 d main of Src.

27.   The peptide of claim 18 the binding affinity of
5 which is at least four-fold greater than that exhibited by
the peptide RPLPPLP (SEQ ID NO:9) for the SH3 domain of Src.

28.   The peptide of claim 1 which further exhibits a
general binding affinity for the SH3 domain of Abl, Grb2,
PLC-δ, PLC-γ, Ras GAP, Nck, p85 PI-3'Kinase, and proteins
10 related thereto.

29.   The peptide of claim 1 which exhibits a selective
binding affinity for the SH3 domain of Src and Src-related
proteins.

30.   The peptide of claim 18 which further exhibits a
15 general binding affinity for the SH3 domain of Abl, Grb2,
PLC-δ, PLC-γ, Ras GAP, Nck, p85 PI-3'Kinase, and proteins
related thereto.

31.   The peptide of claim 18 which exhibits a selective
binding affinity for the SH3 domain of Src and Src-related
20 proteins.

32.   A peptide having the amino acid sequence
RSTPRPLPMLPTTR (SEQ ID NO:62).

33.   A peptide having the amino acid sequence
RSTPRPLPPLPTTR (SEQ ID NO:67).

25   34.   A peptide having the amino acid sequence
GILAPPVPPRNTR (SEQ ID NO:63).

35.   A peptide having the amino acid sequence
VLKRPLPIPPVTR (SEQ ID NO:64).

36.   A peptide having the amino acid sequence
30 GPHRRLPPTPATR (SEQ ID NO:65).

37.   A peptide having the amino acid sequence
ANPSPATRPLPTR (SEQ ID NO:66).

38.   A peptide having an amino acid sequence selected
from the group consisting of RSTRPLPILPRTT, STPRPLPMLPTTR,
35 STNRPLPMIPTTR, RSTRPLPSLPITT, STSRPLPSLPTTR, RSTRSLPPLPPTT,
RSTRQLPIPPTTT, STPRPLPLIPTTP, RSTRPLPPTPLTT, and
RSTRPQPPPPITT (SEQ ID NOS:85-94).

- 92 -

SUBSTITUTE SHEET (RULE 26)

39. A peptid having the amino acid s quenc selected
from the group consisting of VLKRPLPIPPVTR (SEQ ID NO:64),
YSTRPVPPITRPS (SEQ ID NO:76), SHKSRLPPLPTRP (SEQ ID NO:77),
GPHRRLPPTPATR (SEQ ID NO:65), PATRPLPTRPSRT (SEQ ID NO:81),
5 and SGGILAPPVPPRN (SEQ ID NO:84).

40. A peptide having the amino acid sequence selected
from the group consisting of PPNSPLPPLPTHL (SEQ ID NO:72),
TGRGPLPPLPNDS (SEQ ID NO:74), YRFRALPSPPSAS, LAQRQLPPTPGRD
ALQRRLPRTPPPA (SEQ ID NOS:78-80), YSTRPLPSRPSRT, and
10 XPGRILLLLPSEPR (SEQ ID NOS:82-83).

41. A construct comprising a nucleic acid encoding a
peptide of claim 1 or its complement.

42. The construct of claim 41 which is a DNA
polynucleotide.

15   43. The construct of claim 41 which is a RNA
polynucleotide.

44. A construct comprising a nucleic acid encoding a
peptide of claim 18 or its complement.

45. The construct of claim 44 which is a DNA
20 polynucleotide.

46. The construct of claim 44 which is a RNA
polynucleotide.

47. The construct of claim 41 which is a transforming
vector.

25   48. The construct of claim 44 which is a transforming
vector.

49. A host cell transformed with the vector of claim
47.

50. A host cell transformed with the vector of claim
30 48.

51. A conjugate comprising a peptide of claim 1 and a
second molecule.

52. The conjugate of claim 51 in which said second
molecule is selected from the group consisting of an amino
35 acid, a peptide, a protein, a nucl ic acid, a nucl oside, a
glyc sidic r sidu , a label, a drug r a small molecule.

- 93 -

SUBSTITUTE SHEET (RULE 26)

53. A diagnostic kit for the detection of SH3 domains comprising an SH3 domain-binding peptide and a detectable label conjugated to said peptide directly, indirectly or by complexation, said peptide comprising: (i) a core sequence

5 motif of the formula RXLPφφP (SEQ ID NO:71), wherein X represents any amino acid except cysteine and φ represents a hydrophobic amino acid residue, each letter representing the standard one-letter designation for the corresponding amino acid residue; and (ii) two or more additional amino acid

10 residues flanking said core sequence at its C-terminal end, N-terminal end or both.

54. A drug delivery system comprising an SH3 domain-binding peptide and a drug conjugated to said peptide directly, indirectly or by complexation, said peptide

15 comprising: (i) a core sequence motif of the formula RXLPφφP (SEQ ID NO:71), wherein X represents any amino acid except cysteine and φ represents a hydrophobic amino acid residue, each letter representing the standard one-letter designation for the corresponding amino acid residue; and (ii) two or

20 more additional amino acid residues flanking said core sequence at its C-terminal end, N-terminal end or both.

55. The drug delivery system of claim 54 which may be administered parenterally, orally, enterally, topically or by inhalation.

25 56. The drug delivery system of claim 54 which may be administered intranasally, opthalmically or intravaginally.

57. The drug delivery system of claim 54 which is in the form of a solid, gel, liquid or aerosol.

58. A method of modulating the activity of Src or Src-

30 related proteins comprising administering a composition comprising an effective amount of a peptide of claim 1 and a carrier.

59. The method of claim 58 which inhibits the activity of Src or Src-related proteins.

35 60. The method of claim 58 which activates Src or Src-related proteins.

- 94 -

61. A method of identifying a peptide having a region
that binds t an SH3 domain c mprising:

(a) providing an imm bilized target pr tein
comprising an SH3 domain;

5   (b) incubating said immobilized target protein
with an aliquot taken from a random peptide library;

(c) washing unbound peptide from said immobilized
target protein;

(d) recovering the peptide bound to said
10 immobilized target protein; and

(e) determining the primary sequence of the SH3
domain-binding peptide.

62. The method of claim 61 in which said library is a
displayed random peptide library.

15   63. The method of claim 62 in which said library is a
phage-displayed random peptide library.

64. The method of claim 62 in which said library is a
phagemid-displayed random peptide library.

65. The method of claim 61 in which step (c) includes
20 washing unbound phage from said immobilized target protein;
step (d) includes recovering the phage bound to said
immobilized target protein; and step (e) includes determining
the relevant nucleotide sequence of said binding phage
nucleic acid, from which the primary sequence corresponding
25 to the SH3 domain-binding peptide is deduced.

66. A method of identifying a peptide having a region
that binds to an SH3 domain comprising:

(a) providing an immobilized target protein
comprising an SH3 domain;

30   (b) incubating said immobilized target protein
with an aliquot taken from a phage-displayed random peptide
library, which library includes peptides having a random
sequence of ≥ 8 amino acid residues;

(c) washing unbound phage from said immobilized
35 target pr tein;

(d) recovering the phage bound to said immobiliz d
targ t protein; and

- 95 -

(e)   determining the r levant nucl otide sequ nce
f said binding phage nucleic acid and d ducing the primary
s quenc c rr sponding to the SH3 d main-binding peptide.

67.   The method of claim 66 which further comprises

5 amplifying the titer of the recovered phage.

68.   The method of claim 66 which further comprises
repeating the steps of incubation, washing and recovery to
provide SH3 domain-binding peptide-enriched phage.

69.   A pharmaceutical composition comprising an SH3

10 domain-binding peptide and a pharmaceutically acceptable
carrier, said peptide comprising: (i) a 9-mer sequence motif
of the formula RXLPφφPXψ (SEQ ID NO:10), wherein X represents
any amino acid except cysteine, φ represents a hydrophobic
amino acid residue, and wherein ψ is a hydrophilic amino acid

15 residue except cysteine, each letter representing the
standard one-letter designation for the corresponding amino
acid residue; and, optionally, (ii) additional amino acid
residues flanking said 9-mer sequence at its C-terminal end,
N-terminal end or both, up to a total of 45 amino acid

20 residues, including said 9-mer sequence.

70.   The composition of claim 69 in which at least one
additional amino acid flanks said 9-mer sequence.

71.   The composition of claim 69 in which at least two
additional amino acids flank said 9-mer sequence.

25   72.   The composition of claim 69 in which at least three
additional amino acids flank said 9-mer sequence.

73.   A method of disrupting protein tyrosine kinase-
mediated signal transduction pathways comprising
administering an effective amount of a peptide of claim 1.

30   74.   A method of regulating the processing, trafficking
or translation of RNA by administering an effective amount of
a peptide of claim 1.

35

- 96 -

1/14

Pro.Gly
GGC.TCG.AGN.(NNB)$_{18}$.CCA.GGT

GGT.CCA.(NNV)$_{18}$.AGA.TCT.GG

N=A,C,G,T
B=C,G,T
V=A,C,G

FILL IN WITH Taq DNA POLYMERASE

Xho I
GGC.TCG.AGN.(NNB)$_{18}$.CCA.GGT

GGT.CCA.(NNV)$_{18}$.AGA.TCT.GG
Xba I

CLEAVE WITH Xho I AND Xba I

TCG.AGN.(NNB)$_{18}$.CCA.GGT

GGT.CCA.(NNV)$_{18}$.AGA.TC

LIGATE WITH Xho I + Xba I-
CLEAVED M13 m663 VECTOR

ELECTROPORATE INTO E. coli DH5α F'

TSAR-9 LIBRARY
OF pIII-RANDOM SEQUENCE FUSION PROTEINS

. . . S H S | S (R/S) X$_{18}$ P G X$_{18}$ S R | P S R T . . .

SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.1

                                       Gly
tt.ttg.tcg.acN.(NNB)$_{10}$.Ngc.ggt.g
_____

                    cg.cca.cNV.(NNV)$_{10}$.tga.tca.ttt.t
N=A,G,T,C
B=G,T,C | FILL IN WITH Taq DNA POLYMERASE
V=G,A,C   ↓

        Sal I
tt.ttg.tcg.acN.(NNB)$_{10}$.Ngc.ggt.g
_____ ....

                    cg.cca.cNV.(NNV)$_{10}$.tga.tca.ttt.t
                                             Spe I

                        | CLEAVE WITH Sal I + Spe I
                        ↓

    tcg.acN.(NNB)$_{10}$.Ngc.ggt.g
_____

                    cg.cca.cNV.(NNV)$_{10}$.tga.tc

                | LIGATE WITH Xho I + Xba I-
                | CLEAVED M13 m663 VECTOR

                | ELECTROPORATE INTO E. coli DH5α F'
                ↓
                TSAR-12 LIBRARY
    OF pIII-RANDOM SEQUENCE FUSION PROTEINS

. . . S H S | S (S/T) X$_{10}$ΦG∂X$_{10}$ T R | P S R T . . .
            ↑
                                Φ=S,R,G,C, OR W
                                ∂=V,A,D,E, OR G

SIGNAL PEPTIDASE CLEAVAGE SITE

                        FIG.2

Xho I

T GAC GTC TCG AGT TGT (NNK)$_8$ TGT GGA TCT AGA AGG ATC

CCT AGA TCT TCC TAG
Xba I

N=A,C,G, OR T
K=G OR T
M=A OR C

FILL IN WITH DNA POLYMERASE

Xho I

Xba I

CLEAVE WITH Xho I AND Xba I

TCG AGT TGT (NNK)$_8$ TGT GGA T

CA ACA (NNM)$_8$ ACA CCT AGA TC

LIGATE INTO m663 VECTOR TREATED
WITH Xho I, Xba I, AND CALF INTESTINE
ALKALINE PHOSPHATASE

ELECTROPORATE INTO E. coli XL1-B
CELLS

S S C (X)$_8$ C G S R P S R T . .

SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.3

SUBSTITUTE SHEET (RULE 26)

Xba I
___

TCG AGT TGT (NNK)₈ TGT GGA T CT AGA TCC ACA (MNN)₈ ACA AC

CA ACA (NNM)₈ ACA CCT A GA TCT AGG TGT (KNN)₈ TGT TGA GCT

Xho I   Xho I

→

LIGATION OF TWO DOUBLE-STRANDED
OLIGONUCLEOTIDES INTO A HEAD-TO-HEAD
ARRANGEMENT AT THE Xba I SITE

INSERTION INTO m663 VECTOR CLEAVED
BY Xho I AND Xba I. ILLEGITIMATE LIGATION
AT THE Xba I SITE OF THE VECTOR.

ELECTROPORATE INTO E. coli XL1-B CELLS

S S C (X)₈ C G S R S T (X)₈ T T R . . .
↑
SIGNAL PEPTIDASE CLEAVAGE SITE

FIG.4

FIG.5

CLONE	SEQUENCE			FREQUENCY
T9.SRC3.2	SSFDRQDVDYSIAEKMPIRPGF	RELPPLP	PSRASFGGGASRPSR	2
T12.SRC3.4	SINVVVTGSVIARGAQS	RPLPIPP	ETRPSR	1
T12.SRC3.6	SIAPVGLRVAHEGGVLK	RPLPIPP	VTRPSR	1
T9.SRC3.4	SSSGYVVPKRLGDMREYNWHPGLHVPPN	SPLPPLP	THLQSSRPSR	1
T9.SRC3.6	SSRGEGWNIISSRPPFLSNSDPGVSNKLTGR	GPLPPLP	NDSRPSR	2
T12.SRC3.7	STAVSFRFMPGGGGAFYST	RPVPPIT	RPSRT	1
T12.SRC3.5	STAHSLVDWGTFSGVSHKS	RLPPLPT	RPSRT	1
T9.SRC3.7	XPGYARIVSYRF	RALPSPP	SASRPSR	1
T12.SRC3.3	SINDVDWHMMNSGGPH	RRLPPTP	ATRPSR	4
T9.SRC3.5	SSDNVARRVHASELIYTDLSPGILLAQ	RQLPPTP	GRDPSHSRPSR	1
T12.SRC3.1	SSESPLMYNRVGALQSLTSVPGSWHHFALQ	RRLPRTP	PPASRPSR	5
T12.SRC3.2	STRVSHSVPGYVGGANPSPAT	RPLPTRP	SRTVESC	19
T9.SRC3.3	SRYNDLGTRPVSEVIKYDYFPGYSQHVITPDGSYST	RPLPSRP	SRTVESC	2
T9.SRC3.8	XPG	RILLLPS	EPRTFYNYGHDSRPSR	1
T12.SRC3.1	SIMYGVSVLSSGSGGILA	PPVPPRN	TRPSR	25
Consensus		RPLPPLP		

CLONE	SEQUENCE			FREQUENCY
R8C.YES3.6	SSCTEKTVSGVCGSRST	RPLPILP	RTTRPSR	1
R8C.YES3.5	SSCMLPTDGVQCGSRSTP	RPLPMLP	TTRPSR	1
R8C.YES3.3/SRC3.1	SSCDGTGFRLNCGSRSTN	RPLPMJP	TTRPSR	3
R8C.YES3.1/SRC3.2	SSCMQGQAGLKCGSRST	RPLPSLP	ITTRPSR	7
R8C.YES3.7	SSCYREKDTVGCGSRSTS	RPLPSLP	TTRPSR	2
R8C.YES3.2	SSCLFEDGAGTCGSRST	RSLPPLP	PTTRPSR	2
R8C.YES3.4	SSCDHTLGLGVCGSRST	RQLPIPP	TTTRPSR	1
R8C.YES3.10	SSCDTGRIAPGCGSRSTP	RPLPLIP	TTPRSTNLNLTSTTTRPSR	2
R8C.YES3.8	SSCGLDNAAKTCGSRST	RPLPPTP	LTTRPSR	2
R8C.YES3.9	SSCSRAHETEMCGSRST	RPQPPPP	ITTRPSR	2
Consensus		RPLPPLP		

FIG.6

FIG.7

FIG.8

FIG.9

PEPTIDES: (MICROMOLAR CONCENTRATION)		
12.SRC3.1	GILAPPVPPRNTR	
12.SRC3.6	VLKRPLPIPPVTR	
LONG CONSENSUS RST P RPLPPLPTTR		
PI3 KINASE	KISPPTPKPRPPRPLPV	
VINCULIN	LAPPKPPLPEGEV	
12.SRC3.1	GILAPPVPPRNTR I	
12.SRC3.6	VLKRPLPIPPVTR I	
LONG CONSENSUS	RSTPPRPLPPLPTTR I	
PI3 KINASE	KISPPTPKPRPPRPLPV I	

FIG.10

FIG.11A

FIG.11B

FIG.11C

FIG.11D

ANNEAL OLIGONUCLEOTIDES

FILL IN

DIGEST WITH
Xho I AND Xba I

GEL PURIFY
CUT INSERT

CLONE INTO
nBAX VECTOR

EXPRESS BPL

CTG TGC CTC GAG K NNK₆ CCA NNK₂ CCA NNK₆ TCT AGA CGT GTC AGT
                                    AGA TCT GCA CAG TCA

CTG TGC CTC GAG K NNK₆ CCA NNK₂ CCA NNK₆ TCT AGA CGT GTC AGT
GAC ACG GAG CTC M NNM₆ GGT NNM₂ AGA TCT GCA CAG TCA

CTG TGC C   TC GAG K NNK₆ CCA NNK₂ CCA NNK₆ T   CT AGA CGT GTC AGT
GAC ACG GAG CT   C M NNM₆ GGT NNM₂ GGT NNM₆ AGA TC   T GCA CAG TCA

TC GAG K NNK₆ CCA NNK₂ CCA NNK₆ T
C M NNM₆ GGT NNM₂ GGT NNM₆ AGA TC

. . .TC GAG K NNK₆ CCA NNK₂ CCA NNK₆ TCT AGA. . .
. . .AG CTC M NNMG GGT NNM₂ GGT NNM₆ AGA CTC. . .

    S S/R X₆P X₂ P X₆ S R . . . . pIII

PEPTIDE CLEAVAGE
SITE

FIG.12

FIG.13

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.

US CL :435/5, 7.1, 172.1, 320.1; 530/300; 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/5, 7.1, 172.1, 320.1; 530/300; 514/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG, STN
search terms: sh3, src, peptides

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	The Journal of Biological Chemistry, Volume 269, Number 39, issued 30 Spetember 1994, Sparks et al., "Identification and Chracterization of Src SH3 Ligands from Phage-displayed Random Peptide Libraries", pages 23853-23856, see entire article.	1-74
X --- Y	Journal of the American Chemical Society, Volume 115, issued 1993, Chen et al. "Biased Combinatorial Libraries: Novel Ligands for the SH3 Domain of Phosphatidyllinositol 3-Kinase", pages 12591-12592, see entire article, especially Tables I and II.	1-6, 24, 26, 28, 29, 53 ------------ 7-23, 25, 27, 30-51, 53-74

☒ Further documents are listed in the continuation of Box C.   ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
		"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 OCTOBER 1995	01 NOV 1995

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer LORA M. GREEN
Facsimile N . (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Science, Volume 249, issued 28 September 1990, Langer, "New Methods of Drug Discovery", pages 1527-1533, see entire article.	51, 52, 54-57
Y	Science, Volume 257, issued 07 August 1992, Cicchetti et al., "Identificat ion of a Protein That Binds to the SH3 Region of Abl and Is Similar to BCR and GAP-rho"; pages 803-806, see entire article.	7-23, 25, 27, 30-50
Y	Gene, Volume 128, issued 1993, Kay et al., "An M13 Phage Library Displaying Random 38-amino-acid Petides as a Source of Novel Sequences with Affinity to Selected Targets", pages 59-65, see entire article.	7-23, 25, 27, 30-48, 61-68

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

G01N 33/53; C12Q 1/70; C12N 5/10. 7/01; C07K 7/00, 14/00; A61K 38/03

THIS PAGE BLANK (USPTO)